Solutions to Axler's Linear Algebra Done Right

Aayush Bajaj

2026-08-29T13:00:00+10:00

Contents

Solutions to every exercise in the fourth edition of Sheldon Axler's *Linear Algebra Done Right* (Springer UTM, 2024) — 732 exercises across sections 1A–9D. The book is open access at linear.axler.net and is filed at Linear Algebra Done Right (Axler).

1 Vector Spaces

Contents

1.1 Exercises 1A

Problem 1A.1. Show that $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in \mathbf{C}$.

Solution. Commutativity of addition in $\mathbf{C}$ is commutativity in $\mathbf{R}$, one coordinate at a time. Write $\alpha = a + bi$ and $\beta = c + di$ with $a, b, c, d \in \mathbf{R}$; then by the definition of complex addition in 1.1,

$$\begin{aligned}\alpha + \beta &= (a + c) + (b + d)i \\ &= (c + a) + (d + b)i \\ &= \beta + \alpha,\end{aligned}$$

the middle equality because addition in $\mathbf{R}$ is commutative, and the outer two because complex numbers are equal exactly when their real and imaginary parts agree (1.1).

Problem 1A.2. Show that $(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.

Solution. Associativity of addition in $\mathbf{C}$ is associativity in $\mathbf{R}$, one coordinate at a time. Write $\alpha = a+bi$, $\beta = c + di$, $\lambda = e + fi$ with $a, b, c, d, e, f \in \mathbf{R}$; two applications of the definition of complex addition in 1.1 give

$$\begin{aligned}(\alpha + \beta) + \lambda &= \big((a + c) + e\big) + \big((b + d) + f\big)i \\ &= \big(a + (c + e)\big) + \big(b + (d + f)\big)i \\ &= \alpha + (\beta + \lambda),\end{aligned}$$

the middle equality because addition in $\mathbf{R}$ is associative, and the outer two because complex numbers are equal exactly when their real and imaginary parts agree (1.1).

Problem 1A.3. Show that $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.

Solution. Both products expand to the same four real summands in each coordinate. Write $\alpha = a + bi$, $\beta = c+di$, $\lambda = e+fi$ with $a, b, c, d, e, f \in \mathbf{R}$. Two applications of the definition of complex multiplication in 1.1, expanded by the real distributive and associative laws, give

$$\begin{aligned}(\alpha\beta)\lambda &= \big((ac - bd)e - (ad + bc)f\big) \\ &\quad + \big((ac - bd)f + (ad + bc)e\big)i \\ &= (ace - bde - adf - bcf) + (acf - bdf + ade + bce)i\end{aligned}$$

and, in the other grouping,

$$\begin{aligned}\alpha(\beta\lambda) &= \big(a(ce - df) - b(cf + de)\big) \\ &\quad + \big(a(cf + de) + b(ce - df)\big)i \\ &= (ace - adf - bcf - bde) + (acf + ade + bce - bdf)i.\end{aligned}$$

The two real parts are the same four summands in a different order, as are the two imaginary parts, so they agree by commutativity and associativity of addition in $\mathbf{R}$. Hence $(\alpha\beta)\lambda = \alpha(\beta\lambda)$.

Problem 1A.4. Show that $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbf{C}$.

Solution. Both sides expand to the same four real summands in each coordinate. Write $\lambda = e + fi$, $\alpha = a+bi$, $\beta = c+di$ with $a, b, c, d, e, f \in \mathbf{R}$. By the definitions of complex addition and multiplication in 1.1, expanded by the real distributive law,

$$\begin{aligned}\lambda(\alpha+\beta) &= \big(e(a+c)-f(b+d)\big) \\ &\quad + \big(e(b+d)+f(a+c)\big)i \\ &= (ea+ec-fb-fd)+(eb+ed+fa+fc)i,\end{aligned}$$

while $\lambda\alpha = (ea-fb)+(eb+fa)i$ and $\lambda\beta = (ec-fd)+(ed+fc)i$, so

$$\lambda\alpha+\lambda\beta = \big((ea-fb)+(ec-fd)\big)+\big((eb+fa)+(ed+fc)\big)i.$$

Each coordinate is the same sum of four real numbers in a different order, hence equal by commutativity and associativity of addition in $\mathbf{R}$. Thus $\lambda(\alpha+\beta)=\lambda\alpha+\lambda\beta$.

Problem 1A.5. Show that for every $\alpha \in \mathbf{C}$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha+\beta=0$.

Solution. Take $\beta = (-a)+(-b)i$, where $\alpha = a+bi$ with $a,b\in\mathbf{R}$; here 0 means $0+0i$.
Existence: by the definition of complex addition in 1.1,

$$\alpha+\beta = \big(a+(-a)\big)+\big(b+(-b)\big)i = 0+0i = 0.$$

Uniqueness: if $\beta = c+di$ with $c,d\in\mathbf{R}$ satisfies $\alpha+\beta=0$, then $(a+c)+(b+d)i = 0+0i$, so $a+c=0$ and $b+d=0$ since complex numbers are equal exactly when their real and imaginary parts agree (1.1). Uniqueness of additive inverses in $\mathbf{R}$ then forces $c=-a$ and $d=-b$, the β exhibited above.

Problem 1A.6. Show that for every $\alpha\in\mathbf{C}$ with $\alpha\neq 0$, there exists a unique $\beta\in\mathbf{C}$ such that $\alpha\beta = 1$.

Solution. Take

$$\beta = \frac{a}{a^2+b^2}+\frac{-b}{a^2+b^2}\,i,$$

where $\alpha = a+bi$ with $a,b\in\mathbf{R}$; here 1 means $1+0i$. Since $\alpha\neq 0$, at least one of a,b is nonzero, so $a^2+b^2>0$ and the two coefficients are real numbers.
Existence: by the definition of complex multiplication in 1.1,

$$\begin{aligned}\alpha\beta &= \left(a\cdot\frac{a}{a^2+b^2}-b\cdot\frac{-b}{a^2+b^2}\right) \\ &\quad + \left(a\cdot\frac{-b}{a^2+b^2}+b\cdot\frac{a}{a^2+b^2}\right)i \\ &= \frac{a^2+b^2}{a^2+b^2}+\frac{-ab+ab}{a^2+b^2}\,i \\ &= 1+0i = 1.\end{aligned}$$

Uniqueness: if $\beta = c+di$ with $c,d\in\mathbf{R}$ satisfies $\alpha\beta=1$, then comparing real and imaginary parts in $(ac-bd)+(ad+bc)i = 1+0i$ gives

$$ac-bd=1 \qquad\text{and}\qquad ad+bc=0.$$

Taking $a\cdot(\text{first})+b\cdot(\text{second})$ yields $(a^2+b^2)c=a$, and $-b\cdot(\text{first})+a\cdot(\text{second})$ yields $(a^2+b^2)d=-b$. Since $a^2+b^2\neq 0$, this forces

$$c=\frac{a}{a^2+b^2},\qquad d=\frac{-b}{a^2+b^2},$$

the β exhibited above.

Problem 1A.7. Show that
$$\frac{-1+\sqrt{3}\,i}{2}$$
is a cube root of 1 (meaning that its cube equals 1).

Solution. Squaring and multiplying once more returns 1. Put $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}\,i$, which is the given number in the standard form of 1.1. By the definition of complex multiplication in 1.1, using $\sqrt{3}\cdot\sqrt{3} = 3$,

$$\begin{aligned}\omega^2 &= \left(\frac{1}{4} - \frac{3}{4}\right)\\ &\quad + \left(-\frac{\sqrt{3}}{4} - \frac{\sqrt{3}}{4}\right) i = -\frac{1}{2} - \frac{\sqrt{3}}{2}\,i,\\ \omega^3 = (\omega^2)\omega &= \left(\frac{1}{4} + \frac{3}{4}\right)\\ &\quad + \left(-\frac{\sqrt{3}}{4} + \frac{\sqrt{3}}{4}\right) i = 1 + 0i = 1,\end{aligned}$$

the grouping $\omega^3 = (\omega^2)\omega$ being legitimate by the associativity of Exercise 1A.3.

Problem 1A.8. Find two distinct square roots of i.

Solution. Take $z = \pm\frac{\sqrt{2}}{2}(1+i)$. These are distinct, since their difference $\sqrt{2}(1+i)$ is nonzero, and each squares to i:
$$\left(\pm\frac{1+i}{\sqrt{2}}\right)^2 = \frac{(1+i)^2}{2} = \frac{1+2i+i^2}{2} = \frac{2i}{2} = i.$$

Problem 1A.9. Find $x \in \mathbf{R}^4$ such that
$$(4,-3,1,7) + 2x = (5,9,-6,8).$$

Solution.
$$x = \tfrac{1}{2}\big((5,9,-6,8) - (4,-3,1,7)\big) = \left(\tfrac{1}{2},\ 6,\ -\tfrac{7}{2},\ \tfrac{1}{2}\right).$$
Indeed $2x = (1,12,-7,1)$, and coordinatewise addition (1.13) gives $(4,-3,1,7) + (1,12,-7,1) = (5,9,-6,8)$. (Check!)

Problem 1A.10. Explain why there does not exist $\lambda \in \mathbf{C}$ such that
$$\lambda(2-3i,\ 5+4i,\ -6+7i) = (12-5i,\ 7+22i,\ -32-9i).$$

Solution. The first coordinate forces $\lambda = 3+2i$, and that value fails in the third. Since scalar multiplication acts coordinatewise (1.18) and lists are equal exactly when they agree in each coordinate, such a λ would satisfy $\lambda(2-3i) = 12-5i$; as $2-3i \neq 0$ it is invertible (1.3, 1.5), so
$$\lambda = \frac{12-5i}{2-3i} = \frac{(12-5i)(2+3i)}{4+9} = \frac{39+26i}{13} = 3+2i.$$
But then the third coordinate gives
$$(3+2i)(-6+7i) = -18+9i+14i^2 = -32+9i \neq -32-9i,$$
a contradiction. Hence no such λ exists.

Problem 1A.11. Show that $(x+y)+z = x+(y+z)$ for all $x, y, z \in \mathbf{F}^n$.

Solution. Associativity in $\mathbf{F}^n$ is associativity in $\mathbf{F}$, one coordinate at a time; addition in $\mathbf{F}$ is associative by 1.3, since $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$ (1.6). Write $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$, $z = (z_1, \dots, z_n)$ with all entries in $\mathbf{F}$ (1.11). Then

$$\begin{aligned}(x+y)+z &= \big((x_1+y_1)+z_1,\ \dots,\ (x_n+y_n)+z_n\big)\\ &= \big(x_1+(y_1+z_1),\ \dots,\ x_n+(y_n+z_n)\big)\\ &= x+(y+z),\end{aligned}$$

the outer equalities by the definition 1.13 of addition in $\mathbf{F}^n$ and the middle one by associativity in $\mathbf{F}$ applied in each of the n coordinates.

Problem 1A.12. Show that $(ab)x = a(bx)$ for all $x \in \mathbf{F}^n$ and all $a, b \in \mathbf{F}$.

Solution. Associativity in $\mathbf{F}^n$ is associativity of multiplication in $\mathbf{F}$, one coordinate at a time; the latter holds by 1.3, since $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$ (1.6). Write $x = (x_1, \dots, x_n)$ with each $x_k \in \mathbf{F}$ (1.11). Then

$$\begin{aligned}(ab)x &= \big((ab)x_1,\ \dots,\ (ab)x_n\big)\\ &= \big(a(bx_1),\ \dots,\ a(bx_n)\big)\\ &= a(bx_1, \dots, bx_n) = a(bx),\end{aligned}$$

the first, third, and fourth equalities by the definition 1.18 of scalar multiplication in $\mathbf{F}^n$ (the third read from right to left, with scalar a), and the second by associativity in $\mathbf{F}$ applied in each of the n coordinates.

Problem 1A.13. Show that $1x = x$ for all $x \in \mathbf{F}^n$.

Solution. The identity $1x = x$ is the identity $1\lambda = \lambda$ in $\mathbf{F}$, one coordinate at a time; the latter holds by the identities and commutativity statements in 1.3, since $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$ (1.6). Write $x = (x_1, \dots, x_n)$ with each $x_k \in \mathbf{F}$ (1.11). Then by the definition 1.18 of scalar multiplication,

$$1x = (1x_1, \dots, 1x_n) = (x_1, \dots, x_n) = x,$$

the second equality because $1x_k = x_k$ in $\mathbf{F}$ for each k.

Problem 1A.14. Show that $\lambda(x+y) = \lambda x + \lambda y$ for all $\lambda \in \mathbf{F}$ and all $x, y \in \mathbf{F}^n$.

Solution. Distributivity in $\mathbf{F}^n$ is distributivity in $\mathbf{F}$, one coordinate at a time; the latter holds by 1.3, since $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$ (1.6). Write $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$ with all entries in $\mathbf{F}$ (1.11). Then

$$\begin{aligned}\lambda(x+y) &= \big(\lambda(x_1+y_1),\ \dots,\ \lambda(x_n+y_n)\big)\\ &= (\lambda x_1 + \lambda y_1,\ \dots,\ \lambda x_n + \lambda y_n)\\ &= (\lambda x_1, \dots, \lambda x_n) + (\lambda y_1, \dots, \lambda y_n) = \lambda x + \lambda y,\end{aligned}$$

the first equality by the definitions 1.13 and 1.18 of addition and scalar multiplication in $\mathbf{F}^n$, the second by distributivity in $\mathbf{F}$ in each of the n coordinates, and the last two by 1.13 and 1.18 read from right to left.

Problem 1A.15. Show that $(a+b)x = ax + bx$ for all $a, b \in \mathbf{F}$ and all $x \in \mathbf{F}^n$.

Solution. Distributivity in $\mathbf{F}^n$ is distributivity in $\mathbf{F}$, one coordinate at a time. Write $x = (x_1, \ldots, x_n)$ with each $x_j \in \mathbf{F}$ (1.11). Then

$$\begin{aligned}(a+b)x &= \big((a+b)x_1, \ldots, (a+b)x_n\big)\\ &= (ax_1 + bx_1, \ldots, ax_n + bx_n)\\ &= (ax_1, \ldots, ax_n) + (bx_1, \ldots, bx_n) = ax + bx,\end{aligned}$$

the first and last equalities by the definitions 1.18 and 1.13 of scalar multiplication and addition in $\mathbf{F}^n$. The second equality is the coordinatewise identity $(a+b)x_j = ax_j + bx_j$: by 1.6 the field $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$, and 1.3 states distributivity in the form $\lambda(\alpha+\beta) = \lambda\alpha + \lambda\beta$, so combining it with the commutativity of multiplication in 1.3 gives

$$(a+b)x_j = x_j(a+b) = x_j a + x_j b = ax_j + bx_j.$$

1.2 Exercises 1B

Problem 1B.1. Prove that $-(-v) = v$ for every $v \in V$.

Solution. Both v and $-(-v)$ are additive inverses of $-v$, and additive inverses are unique (1.27). Indeed $v + (-v) = 0$ by the definition 1.28 of $-v$, so $(-v) + v = 0$ by commutativity of addition, which says exactly that v is an additive inverse of $-v$; and $-(-v)$ is one by definition. Hence $-(-v) = v$.

Problem 1B.2. Suppose $a \in \mathbb{F}$, $v \in V$, and $av = 0$. Prove that $a = 0$ or $v = 0$.

Solution. If $a \neq 0$ then $v = 0$, which gives the disjunction. Indeed $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$ (1.1), so $a \neq 0$ has an inverse $\frac{1}{a} \in \mathbb{F}$, and

$$v = 1v = \left(\frac{1}{a}\,a\right)v = \frac{1}{a}(av) = \frac{1}{a}\,0 = 0,$$

using in turn the multiplicative identity property of 1.20, $\frac{1}{a}a = 1$ in $\mathbb{F}$, the associativity property $(bc)v = b(cv)$ of 1.20, the hypothesis $av = 0$, and 1.31.

Problem 1B.3. Suppose $v, w \in V$. Explain why there exists a unique $x \in V$ such that

$$v + 3x = w.$$

Solution. The unique solution is $x = \frac{1}{3}(w - v) = \frac{1}{3}\big(w + (-v)\big)$, which lies in V because $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, so $\frac{1}{3} \in \mathbb{F}$.
Existence: by associativity of scalar multiplication and the multiplicative identity property of 1.20, $3x = (3 \cdot \frac{1}{3})(w - v) = w + (-v)$, so by commutativity and associativity of addition,

$$v + 3x = v + \big(w + (-v)\big) = w + \big(v + (-v)\big) = w + 0 = w.$$

Uniqueness: if $v + 3x = w = v + 3x'$, then adding $-v$ on the left and using associativity, $(-v) + v = 0$, and the additive identity property gives $3x = 3x'$; multiplying by $\frac{1}{3}$ and using associativity of scalar multiplication with the multiplicative identity property gives

$$x = \left(\tfrac{1}{3} \cdot 3\right)x = \tfrac{1}{3}(3x) = \tfrac{1}{3}(3x') = x'.$$

Problem 1B.4. The empty set is not a vector space. The empty set fails to satisfy only one of the requirements listed in the definition of a vector space (1.20). Which one?

Solution. The **additive identity** condition: there is no $0 \in \emptyset$, since $\emptyset$ has no elements.
It is the only failure because it is the only requirement of 1.20 asserting the existence of an element of V; the other five (commutativity, the two associativities, additive inverse, multiplicative identity, and the two distributive properties) are all universally quantified over elements of V, hence vacuously true when $V = \emptyset$. The operations themselves are legitimate: $\emptyset \times \emptyset = \emptyset$ and $\mathbb{F} \times \emptyset = \emptyset$, and the empty function serves as each in the sense of 1.19.

Problem 1B.5. Show that in the definition of a vector space (1.20), the additive inverse condition can be replaced with the condition that

$$0v = 0 \quad \text{for all } v \in V.$$

Here the 0 on the left side is the number 0, and the 0 on the right side is the additive identity of V. [The phrase a "condition can be replaced" in a definition means that the collection of objects satisfying the definition is unchanged if the original condition is replaced with the new condition.]

Solution. Granted the other five conditions of 1.20, the additive inverse condition and the condition $0v = 0$ each imply the other, so the two definitions have the same objects. (The phrase "the additive identity of V" is unambiguous in both settings: the proof of 1.26 uses only the additive identity condition and commutativity.)
(i) If V also satisfies the additive inverse condition, it is a vector space, so $0v = 0$ for every $v \in V$ by 1.30.
(ii) If instead V satisfies $0v = 0$ for all $v \in V$, then for $v \in V$ the vector $(-1)v \in V$ is an additive inverse of v:

$$v + (-1)v = 1v + (-1)v = \big(1 + (-1)\big)v = 0v = 0,$$

using the multiplicative identity condition, the distributive property $(a + b)v = av + bv$, $1 + (-1) = 0$ in $\mathbb{F}$, and the assumed condition.

Problem 1B.6. Let ∞ and $-\infty$ denote two distinct objects, neither of which is in $\mathbb{R}$. Define an addition and scalar multiplication on $\mathbb{R} \cup \{\infty, -\infty\}$ as you could guess from the notation. Specifically, the sum and product of two real numbers is as usual, and for $t \in \mathbb{R}$ define

$$t\infty = \begin{cases} -\infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ \infty & \text{if } t > 0, \end{cases} \qquad t(-\infty) = \begin{cases} \infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ -\infty & \text{if } t > 0, \end{cases}$$

and

$$\begin{aligned} t + \infty = \infty + t &= \infty + \infty = \infty, \\ t + (-\infty) = (-\infty) + t &= (-\infty) + (-\infty) = -\infty, \\ \infty + (-\infty) &= (-\infty) + \infty = 0. \end{aligned}$$

With these operations of addition and scalar multiplication, is $\mathbb{R} \cup \{\infty, -\infty\}$ a vector space over $\mathbb{R}$? Explain.

Solution. No: addition on $W = \mathbb{R} \cup \{\infty, -\infty\}$ is not associative. Take $u = -\infty$, $v = \infty$, and $w = 1$. Then

$$(u + v) + w = \big((-\infty) + \infty\big) + 1 = 0 + 1 = 1,$$

while

$$u + (v + w) = (-\infty) + (\infty + 1) = (-\infty) + \infty = 0,$$

using $\infty + 1 = \infty$ from the definition. Since $1 \neq 0$, the associativity requirement of 1.20 fails, so W is not a vector space over $\mathbb{R}$.

The distributive property $(a+b)v = av+bv$ fails too: with $a=2$, $b=-1$, $v=\infty$,

$$(a+b)v = 1\cdot\infty = \infty \quad \text{while} \quad av+bv = \infty + (-\infty) = 0.$$

Problem 1B.7. Suppose S is a nonempty set. Let V^S denote the set of functions from S to V. Define a natural addition and scalar multiplication on V^S, and show that V^S is a vector space with these definitions.

Solution. Define the operations pointwise, imitating 1.24: for $f, g \in V^S$ and $\lambda \in \mathbb{F}$,

$$(f+g)(x) = f(x)+g(x), \qquad (\lambda f)(x) = \lambda\big(f(x)\big) \qquad \text{for all } x \in S.$$

Both right sides lie in V, so these are an addition and a scalar multiplication on V^S in the sense of 1.19.

Each requirement of 1.20 then holds because two functions $S \to V$ are equal exactly when they agree at every $x \in S$, and at each such x the requirement is the corresponding property of V. Commutativity is typical:

$$(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x),$$

and associativity of addition, associativity and the multiplicative identity for scalar multiplication, and both distributive properties go the same way. (Check!) For the two existence requirements, take $\mathbf{0} \in V^S$ with $\mathbf{0}(x) = 0$ and $-f \in V^S$ with $(-f)(x) = -\big(f(x)\big)$ (well defined by 1.27); then $(f+\mathbf{0})(x) = f(x)+0 = f(x)$ and $\big(f+(-f)\big)(x) = 0 = \mathbf{0}(x)$ for every $x \in S$, so $f+\mathbf{0} = f$ and $f+(-f) = \mathbf{0}$.

Hence V^S is a vector space over $\mathbb{F}$.

Problem 1B.8. Suppose V is a real vector space.

- The *complexification* of V, denoted by $V_{\mathbf{C}}$, equals $V \times V$. An element of $V_{\mathbf{C}}$ is an ordered pair (u,v), where $u, v \in V$, but we write this as $u+iv$.
- Addition on $V_{\mathbf{C}}$ is defined by

$$(u_1+iv_1)+(u_2+iv_2) = (u_1+u_2)+i(v_1+v_2)$$

for all $u_1, v_1, u_2, v_2 \in V$.

- Complex scalar multiplication on $V_{\mathbf{C}}$ is defined by

$$(a+bi)(u+iv) = (au-bv)+i(av+bu)$$

for all $a, b \in \mathbf{R}$ and all $u, v \in V$.

Prove that with the definitions of addition and scalar multiplication as above, $V_{\mathbf{C}}$ is a complex vector space.

[Think of V as a subset of $V_{\mathbf{C}}$ by identifying $u \in V$ with $u+i0$. The construction of $V_{\mathbf{C}}$ from V can then be thought of as generalizing the construction of $\mathbf{C}^n$ from $\mathbf{R}^n$.]

Solution. Each requirement of 1.20 for $V_{\mathbf{C}}$ over $\mathbf{C}$ reduces to the real vector space properties of V applied to the two components of $u+iv$, which is by definition the ordered pair (u,v); so $u_1+iv_1 = u_2+iv_2$ exactly when $u_1 = u_2$ and $v_1 = v_2$. Both formulas return elements of $V_{\mathbf{C}}$, and the representation $a+bi$ with $a, b \in \mathbf{R}$ of a complex number is unique, so the operations are an addition and a scalar multiplication in the sense of 1.19. Throughout, $\alpha u - \beta v$ means $\alpha u + (-\beta)v$: by 1.28 it means $\alpha u + \big(-(\beta v)\big)$, and $\beta v + (-\beta)v = \big(\beta+(-\beta)\big)v = 0v = 0$ by distributivity in V and 1.30, so $-(\beta v) = (-\beta)v$ by uniqueness of additive inverses (1.27).

Fix $u, v, u_1, v_1, u_2, v_2 \in V$ and $\lambda = a+bi$, $\mu = c+di$ in $\mathbf{C}$, with $a, b, c, d \in \mathbf{R}$.

Commutativity and associativity of addition hold componentwise from those properties in V. (Check!) The vector $\mathbf{0} = 0+i0$ is an additive identity, since $(u+iv)+(0+i0) = (u+0)+i(v+0) = u+iv$, and $(-u)+i(-v)$ is an additive inverse of $u+iv$ by the same computation. The multiplicative identity condition holds because $1 = 1 + 0i$:

$$1(u + iv) = (1u - 0v) + i(1v + 0u) = u + iv,$$

using $1u = u$ in V and $0u = 0v = 0$ (1.30).
Associativity of scalar multiplication: $\lambda\mu = (ac - bd) + (ad + bc)i$ gives

$$\begin{aligned}(\lambda\mu)(u + iv) = \big((ac - bd)u - (ad + bc)v\big)\\ + i\big((ac - bd)v + (ad + bc)u\big),\end{aligned}$$

while $\mu(u + iv) = (cu - dv) + i(cv + du)$ gives

$$\begin{aligned}\lambda\big(\mu(u + iv)\big) &= \big(a(cu - dv) - b(cv + du)\big)\\ &\quad + i\big(a(cv + du) + b(cu - dv)\big)\\ &= \big((ac)u - (ad)v - (bc)v - (bd)u\big)\\ &\quad + i\big((ac)v + (ad)u + (bc)u - (bd)v\big),\end{aligned}$$

by the distributive property $\alpha(x+y) = \alpha x+\alpha y$ and associativity $\alpha(\beta u) = (\alpha\beta)u$ in V; the two results agree by $(\alpha + \beta)x = \alpha x + \beta x$ in V with commutativity and associativity of addition.
First distributive property: with $x = u_1 + iv_1$ and $y = u_2 + iv_2$,

$$\begin{aligned}\lambda(x + y) &= \big(a(u_1 + u_2) - b(v_1 + v_2)\big)\\ &\quad + i\big(a(v_1 + v_2) + b(u_1 + u_2)\big)\\ &= \big((au_1 - bv_1) + (au_2 - bv_2)\big)\\ &\quad + i\big((av_1 + bu_1) + (av_2 + bu_2)\big) = \lambda x + \lambda y,\end{aligned}$$

the second equality by $\alpha(x + y) = \alpha x + \alpha y$ in V (with the real scalars a and $-b$) and regrouping by commutativity and associativity of addition in V.
Second distributive property: since $\lambda + \mu = (a + c) + (b + d)i$, with $x = u + iv$,

$$\begin{aligned}(\lambda + \mu)x &= \big((a + c)u - (b + d)v\big) + i\big((a + c)v + (b + d)u\big)\\ &= \big((au - bv) + (cu - dv)\big)\\ &\quad + i\big((av + bu) + (cv + du)\big) = \lambda x + \mu x,\end{aligned}$$

the second equality by $(\alpha + \beta)x = \alpha x + \beta x$ in V (with the real scalars a, c and $-b, -d$, noting $-(b + d) = (-b) + (-d)$) and the same regrouping.
Hence $V_{\mathbf{C}}$ is a complex vector space.

1.3 Exercises 1C

Problem 1C.1. For each of the following subsets of $\mathbf{F}^3$, determine whether it is a subspace of $\mathbf{F}^3$.
(a) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$
(b) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$
(c) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1x_2x_3 = 0\}$
(d) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}$

Solution. Subspaces in (a) and (d), not in (b) and (c). Throughout, U is a subspace exactly when $0 \in U$ and U is closed under addition and under scalar multiplication (1.34).
(a) Since $0 + 2 \cdot 0 + 3 \cdot 0 = 0$, the vector $(0, 0, 0)$ lies in U; and for $x, y \in U$ and $\lambda \in \mathbf{F}$,

$$\begin{aligned}(x_1 + y_1) + 2(x_2 + y_2) + 3(x_3 + y_3) &= (x_1 + 2x_2 + 3x_3) + (y_1 + 2y_2 + 3y_3)\\ &= 0 + 0 = 0,\end{aligned}$$

while $\lambda x_1 + 2\lambda x_2 + 3\lambda x_3 = \lambda(x_1 + 2x_2 + 3x_3) = 0$. So $x + y$ and λx lie in U.

(b) The additive identity is missing: $0 + 2 \cdot 0 + 3 \cdot 0 = 0 \neq 4$.
(c) Closure under addition fails: $(1,1,0)$ and $(0,0,1)$ lie in the set, but their sum $(1,1,1)$ does not, since $1 \cdot 1 \cdot 1 = 1 \neq 0$.
(d) Since $0 = 5 \cdot 0$, the vector $(0,0,0)$ lies in U; and if $x_1 = 5x_3$ and $y_1 = 5y_3$, then $x_1 + y_1 = 5(x_3 + y_3)$ and $\lambda x_1 = \lambda(5x_3) = 5(\lambda x_3)$, which say exactly that $x + y \in U$ and $\lambda x \in U$.

Problem 1C.2. Verify all assertions about subspaces in Example 1.35.

Solution. All five assertions hold; each is the criterion 1.34 applied to the defining condition, which in every case is preserved by sums and scalar multiples and satisfied by the zero vector.
(a) $U = \{x \in \mathbf{F}^4 : x_3 = 5x_4 + b\}$ is a subspace if and only if $b = 0$. For $b \neq 0$ the zero vector fails the condition, since $0 = 5 \cdot 0 + b$ forces $b = 0$. For $b = 0$ it satisfies it, and $x_3 = 5x_4$, $y_3 = 5y_4$ give $x_3 + y_3 = 5(x_4 + y_4)$ and $\lambda x_3 = 5(\lambda x_4)$.
(b) $C[0,1]$ is a subspace of $\mathbf{R}^{[0,1]}$: the zero function is constant, hence continuous, and sums and scalar multiples of continuous functions are continuous.
(c) The set D of differentiable functions on $\mathbf{R}$ is a subspace of $\mathbf{R}^{\mathbf{R}}$: the zero function is differentiable, and $f + g$ and λf are differentiable with $(f+g)' = f' + g'$ and $(\lambda f)' = \lambda f'$.
(d) $U = \{f \in \mathbf{R}^{(0,3)} : f \text{ is differentiable and } f'(2) = b\}$ is a subspace if and only if $b = 0$. The zero function has derivative 0 at 2, so it lies in U exactly when $b = 0$; and when $b = 0$, the derivative rules of (c) give $(f+g)'(2) = 0 + 0 = 0$ and $(\lambda f)'(2) = \lambda \cdot 0 = 0$.
(e) The sequences in $\mathbf{C}^\infty$ with limit 0 form a subspace: $(0,0,\ldots)$ has limit 0, and the limit laws give $\lim_{n\to\infty}(x_n + y_n) = 0$ and $\lim_{n\to\infty} \lambda x_n = 0$ whenever $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = 0$.

Problem 1C.3. Show that the set of differentiable real-valued functions f on the interval $(-4,4)$ such that $f'(-1) = 3f(2)$ is a subspace of $\mathbf{R}^{(-4,4)}$.

Solution. The set U in question satisfies the three conditions of 1.34. The zero function is differentiable with $0'(-1) = 0 = 3 \cdot 0(2)$, so $0 \in U$; and for $f, g \in U$ and $\lambda \in \mathbf{R}$, sums and scalar multiples of differentiable functions are differentiable, with $(f+g)' = f' + g'$ and $(\lambda f)' = \lambda f'$, so

$$(f+g)'(-1) = 3f(2) + 3g(2) = 3(f+g)(2), \qquad (\lambda f)'(-1) = \lambda \cdot 3f(2) = 3(\lambda f)(2).$$

Hence $f + g \in U$ and $\lambda f \in U$, and by 1.34 U is a subspace of $\mathbf{R}^{(-4,4)}$.

Problem 1C.4. Suppose $b \in \mathbf{R}$. Show that the set of continuous real-valued functions f on the interval $[0,1]$ such that $\int_0^1 f = b$ is a subspace of $\mathbf{R}^{[0,1]}$ if and only if $b = 0$.

Solution. The set U in question is a subspace exactly when $b = 0$.
(i) $b \neq 0$: then $\int_0^1 0 = 0 \neq b$, so U omits the additive identity and by 1.34 is not a subspace.
(ii) $b = 0$: then $0 \in U$, and for $f, g \in U$ and $\lambda \in \mathbf{R}$ the functions $f + g$ and λf are continuous, with

$$\int_0^1 (f+g) = \int_0^1 f + \int_0^1 g = 0, \qquad \int_0^1 (\lambda f) = \lambda \int_0^1 f = 0$$

by linearity of the integral. So $f + g, \lambda f \in U$ and 1.34 makes U a subspace of $\mathbf{R}^{[0,1]}$.

Problem 1C.5. Is $\mathbf{R}^2$ a subspace of the complex vector space $\mathbf{C}^2$?

Solution. No: the scalars of $\mathbf{C}^2$ are all of $\mathbf{C}$, and $(1,0) \in \mathbf{R}^2$ with $i \in \mathbf{C}$ gives

$$i(1,0) = (i,0) \notin \mathbf{R}^2.$$

Closure under scalar multiplication fails, so by 1.34 $\mathbf{R}^2$ is not a subspace of the complex vector space $\mathbf{C}^2$.

Problem 1C.6. (a) Is $\{(a,b,c) \in \mathbf{R}^3 : a^3 = b^3\}$ a subspace of $\mathbf{R}^3$?
(b) Is $\{(a,b,c) \in \mathbf{C}^3 : a^3 = b^3\}$ a subspace of $\mathbf{C}^3$?

Solution. (a) Yes, because cubing is injective on $\mathbf{R}$, so the set is $\{(a,a,c) : a,c \in \mathbf{R}\}$. Indeed, if $a,b \in \mathbf{R}$ with $a^3 = b^3$ then

$$0 = a^3 - b^3 = (a-b)(a^2+ab+b^2), \qquad a^2+ab+b^2 = \left(a + \tfrac{b}{2}\right)^2 + \tfrac{3}{4}b^2,$$

and the second factor vanishes only when $a = b = 0$; otherwise $a - b = 0$. Either way $a = b$. The set $\{(a,a,c)\}$ contains 0 and is closed under addition and scalar multiplication, since both operations act coordinatewise and preserve equality of the first two coordinates (Check!), so 1.34 makes it a subspace of $\mathbf{R}^3$.
(b) No: cubing is not injective on $\mathbf{C}$. With $\omega = e^{2\pi i/3}$, so $\omega^3 = 1$ and $\omega + \omega^2 = -1$, the vectors $u = (1,1,0)$ and $v = (1,\omega,0)$ lie in the set, while their sum $(2, 1+\omega, 0)$ does not, because

$$(1+\omega)^3 = 1 + 3(\omega+\omega^2) + \omega^3 = 2 - 3 = -1 \neq 8 = 2^3.$$

Closure under addition fails, so by 1.34 the set is not a subspace of $\mathbf{C}^3$.

Problem 1C.7. Prove or give a counterexample: If U is a nonempty subset of $\mathbf{R}^2$ such that U is closed under addition and under taking additive inverses (meaning $-u \in U$ whenever $u \in U$), then U is a subspace of $\mathbf{R}^2$.

Solution. False. Take $U = \mathbf{Z}^2 = \{(m,n) \in \mathbf{R}^2 : m,n \in \mathbf{Z}\}$, which is nonempty and closed under addition and negation because $\mathbf{Z}$ is (Check!). But

$$\tfrac{1}{2}(1,0) = \left(\tfrac{1}{2}, 0\right) \notin U,$$

so closure under scalar multiplication fails and by 1.34 U is not a subspace of $\mathbf{R}^2$.

Problem 1C.8. Give an example of a nonempty subset U of $\mathbb{R}^2$ such that U is closed under scalar multiplication, but U is not a subspace of $\mathbb{R}^2$.

Solution. Take the union of the two coordinate axes,

$$U = \{(x_1,x_2) \in \mathbb{R}^2 : x_1x_2 = 0\},$$

which is nonempty and closed under scalar multiplication because $(\lambda x_1)(\lambda x_2) = \lambda^2(x_1x_2) = 0$. It is not closed under addition, since $(1,0),(0,1) \in U$ but $(1,0)+(0,1) = (1,1) \notin U$. By 1.34, U is not a subspace of $\mathbb{R}^2$.

Problem 1C.9. A function $f\colon \mathbb{R} \to \mathbb{R}$ is called periodic if there exists a positive number p such that $f(x) = f(x+p)$ for all $x \in \mathbb{R}$. Is the set of periodic functions from $\mathbb{R}$ to $\mathbb{R}$ a subspace of $\mathbb{R}^{\mathbb{R}}$? Explain.

Solution. No: the set P of periodic functions is not closed under addition. Take

$$f(x) = \cos x, \qquad g(x) = \cos(\sqrt{2}\,x), \qquad h = f + g,$$

with periods 2π and $\sqrt{2}\,\pi$, so $f,g \in P$.
Since $\cos \leq 1$ we have $h \leq 2$, and $h(0) = 2$; moreover $h(x) = 2$ forces both $\cos x = 1$ and $\cos(\sqrt{2}\,x) = 1$, that is $x = 2\pi m$ and $\sqrt{2}\,x = 2\pi n$ with $m,n \in \mathbb{Z}$. Substituting gives $\sqrt{2}\,m = n$, which forces $m = 0$ (else $\sqrt{2} = n/m$ would be rational) and hence $x = 0$. So

$$\{x \in \mathbb{R} : h(x) = 2\} = \{0\}.$$

If h had period $p > 0$ then $h(p) = h(0) = 2$, forcing $p = 0$, a contradiction. Thus $f, g \in P$ but $f + g \notin P$, and by 1.34 P is not a subspace of $\mathbb{R}^{\mathbb{R}}$.

Problem 1C.10. Suppose V_1 and V_2 are subspaces of V. Prove that the intersection $V_1 \cap V_2$ is a subspace of V.

Solution. Each of the three conditions of 1.34 passes to the intersection because it holds in V_1 and in V_2 separately.
Indeed $0 \in V_1$ and $0 \in V_2$, so $0 \in V_1 \cap V_2$; and if $u, w \in V_1 \cap V_2$ and $a \in \mathbb{F}$, then $u + w$ and au lie in V_1 (which is closed under both operations) and likewise in V_2, hence in $V_1 \cap V_2$. By 1.34, $V_1 \cap V_2$ is a subspace of V.

Problem 1C.11. Prove that the intersection of every collection of subspaces of V is a subspace of V.

Solution. The argument of Exercise 10 runs verbatim with an arbitrary index set, since each condition of 1.34 need only be checked one member at a time. Let $\{V_i\}_{i \in I}$ be subspaces of V with $I \neq \varnothing$ and put $W = \bigcap_{i \in I} V_i$.
Every V_i contains 0, so $0 \in W$. If $u, w \in W$ and $a \in \mathbb{F}$, then for each $i \in I$ we have $u, w \in V_i$, whence $u + w \in V_i$ and $au \in V_i$ by closure in V_i; as i was arbitrary, $u + w, au \in W$. By 1.34, W is a subspace of V.

Problem 1C.12. Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.

Solution. Let U, W be subspaces of V.
If $W \subseteq U$ (or symmetrically $U \subseteq W$), then $U \cup W = U$, a subspace.
Conversely suppose $U \cup W$ is a subspace and neither inclusion holds, so there are $u \in U \setminus W$ and $w \in W \setminus U$. Closure under addition gives $u + w \in U \cup W$, and both alternatives fail:
(i) $u + w \in U$ forces $w = (u + w) - u \in U$, contradicting $w \notin U$;
(ii) $u + w \in W$ forces $u = (u + w) - w \in W$, contradicting $u \notin W$;
each using that U, respectively W, is closed under subtraction. Hence one of U, W contains the other.

Problem 1C.13. Prove that the union of three subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.
This exercise is surprisingly harder than Exercise 12, possibly because this exercise is not true if we replace $\mathbb{F}$ with a field containing only two elements.

Solution. Let V_1, V_2, V_3 be subspaces of V and $W = V_1 \cup V_2 \cup V_3$. If, say, $V_2 \cup V_3 \subseteq V_1$, then $W = V_1$ is a subspace.
Conversely suppose W is a subspace but no V_j contains the other two.
First, no V_j lies in the union of the other two: if say $V_3 \subseteq V_1 \cup V_2$, then $W = V_1 \cup V_2$ is a subspace, so by Exercise 12 one contains the other, say $V_2 \subseteq V_1$, whence $W = V_1$ and V_1 contains V_2 and V_3, contrary to assumption. So we may choose

$$u \in V_1 \setminus (V_2 \cup V_3), \qquad w \in V_2 \setminus (V_1 \cup V_3).$$

Next, $u + \lambda w \in V_3$ for every nonzero $\lambda \in \mathbb{F}$. Indeed $u + \lambda w \in W$, and
(i) $u + \lambda w \in V_1$ would give $\lambda w = (u + \lambda w) - u \in V_1$, hence $w = \lambda^{-1}(\lambda w) \in V_1$, contradicting the choice of w;
(ii) $u + \lambda w \in V_2$ would give $u = (u + \lambda w) - \lambda w \in V_2$ since $\lambda w \in V_2$, contradicting the choice of u.
Finally, $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, so 1 and 2 are distinct nonzero scalars; applying the previous paragraph to each

and subtracting inside the subspace V_3,

$$w = (u + 2w) - (u + w) \in V_3,$$

contradicting $w \notin V_3$. Hence one of V_1, V_2, V_3 contains the other two.

Problem 1C.14. Suppose

$$U = \{(x, -x, 2x) \in \mathbb{F}^3 : x \in \mathbb{F}\} \qquad \text{and} \qquad W = \{(x, x, 2x) \in \mathbb{F}^3 : x \in \mathbb{F}\}.$$

Describe $U + W$ using symbols, and also give a description of $U + W$ that uses no symbols.

Solution. We have

$$U + W = \{(x, y, 2x) \in \mathbb{F}^3 : x, y \in \mathbb{F}\},$$

that is, $U + W$ is the set of vectors in $\mathbb{F}^3$ whose third coordinate is twice the first.
By 1.36 the general element of $U + W$ is

$$(a, -a, 2a) + (b, b, 2b) = \big(a + b,\ b - a,\ 2(a + b)\big),$$

which has the stated form with $x = a + b$, $y = b - a$; this gives the inclusion $\subseteq$. Conversely, given $x, y \in \mathbb{F}$, the scalar 2 is invertible in $\mathbb{F}$, so $a = (x - y)/2$ and $b = (x + y)/2$ satisfy $a + b = x$ and $b - a = y$, and the display exhibits $(x, y, 2x)$ in $U + W$.

Problem 1C.15. Suppose U is a subspace of V. What is $U + U$?

Solution. $U + U = U$. By 1.40 the sum $U + U$ is the smallest subspace of V containing U and U, and that is U itself, since U is already a subspace.
Method (2): by 1.36, $U + U = \{u_1 + u_2 : u_1, u_2 \in U\} \subseteq U$ since U is closed under addition, while $u = u + 0 \in U + U$ for each $u \in U$ since $0 \in U$.

Problem 1C.16. Is the operation of addition on the subspaces of V commutative? In other words, if U and W are subspaces of V, is $U + W = W + U$?

Solution. Yes: $U + W = W + U$, because addition in V is commutative (an axiom in 1.20). Indeed, by 1.36 each $v \in U + W$ has the form

$$v = u + w = w + u \qquad (u \in U,\ w \in W),$$

so $v \in W + U$; hence $U + W \subseteq W + U$, and interchanging U and W gives the reverse inclusion.

Problem 1C.17. Is the operation of addition on the subspaces of V associative? In other words, if V_1, V_2, V_3 are subspaces of V, is

$$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)?$$

Solution. Yes: both sides equal

$$S = \{v_1 + v_2 + v_3 : v_1 \in V_1,\ v_2 \in V_2,\ v_3 \in V_3\},$$

the set $V_1 + V_2 + V_3$ of 1.36, unambiguous because addition in V is associative (1.20).
Since $V_1 + V_2$ is a subspace (1.40), applying 1.36 twice writes each $v \in (V_1 + V_2) + V_3$ as $v = (v_1 + v_2) + v_3 = v_1 + v_2 + v_3 \in S$ with $v_k \in V_k$; conversely each such $v_1 + v_2 + v_3$ equals $(v_1 + v_2) + v_3 \in (V_1 + V_2) + V_3$. So $(V_1 + V_2) + V_3 = S$, and moving the parentheses gives $V_1 + (V_2 + V_3) = S$ by the identical argument.

Problem 1C.18. Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?

Solution. Yes: $\{0\}$ is the unique additive identity, and $\{0\}$ is the only subspace with an additive inverse. For any subspace U, every element of $U+\{0\}$ is $u+0=u\in U$, and every $u\in U$ is $u+0\in U+\{0\}$, so

$$U+\{0\}=\{0\}+U=U$$

using 1C.16 for the second equality; thus $\{0\}$ is a two-sided identity. If E is another identity, then $\{0\}+E=\{0\}$ by taking $U=\{0\}$ in its defining property, while the display with $U=E$ gives $\{0\}+E=E$; hence $E=\{0\}$.
For inverses, suppose $U+W=\{0\}$. Each $u\in U$ equals $u+0\in U+W$, so $U\subseteq\{0\}$ and $U=\{0\}$; likewise $W=\{0\}$. Conversely $\{0\}+\{0\}=\{0\}$, so $\{0\}$ is its own inverse.

Problem 1C.19. Prove or give a counterexample: If V_1, V_2, U are subspaces of V such that

$$V_1+U=V_2+U,$$

then $V_1=V_2$.

Solution. False. In $V=\mathbf{F}^2$ take

$$V_1=\{(x,0):x\in\mathbf{F}\},\qquad V_2=\{(0,y):y\in\mathbf{F}\},\qquad U=\mathbf{F}^2,$$

each a subspace by 1.34 (Check!). Since $V_k\subseteq U$, by 1.40 the smallest subspace containing V_k and U is U itself, so

$$V_1+U=\mathbf{F}^2=V_2+U.$$

But $(1,0)\in V_1\setminus V_2$, so $V_1\neq V_2$.

Problem 1C.20. Suppose

$$U=\{(x,x,y,y)\in\mathbf{F}^4:x,y\in\mathbf{F}\}.$$

Find a subspace W of $\mathbf{F}^4$ such that $\mathbf{F}^4=U\oplus W$.

Solution. Take

$$W=\{(x,0,y,0)\in\mathbf{F}^4:x,y\in\mathbf{F}\},$$

a subspace of $\mathbf{F}^4$ by 1.34, since it contains 0 and both operations act coordinatewise, preserving vanishing of the second and fourth coordinates (Check!).
The sum is all of $\mathbf{F}^4$: for $(a,b,c,d)\in\mathbf{F}^4$,

$$(a,b,c,d)=(b,b,d,d)+(a-b,0,c-d,0)\in U+W.$$

It is direct by 1.46, because $v\in U\cap W$ has the form $v=(x,x,y,y)$ with second and fourth coordinates 0, forcing $x=y=0$ and $v=0$. Hence $\mathbf{F}^4=U\oplus W$.

Problem 1C.21. Suppose

$$U=\{(x,y,x+y,x-y,2x)\in\mathbf{F}^5:x,y\in\mathbf{F}\}.$$

Find a subspace W of $\mathbf{F}^5$ such that $\mathbf{F}^5=U\oplus W$.

Solution. Take

$$W=\{(0,0,z_1,z_2,z_3)\in\mathbf{F}^5:z_1,z_2,z_3\in\mathbf{F}\}.$$

Here W is a subspace of $\mathbf{F}^5$ by 1.34: it contains 0, and both operations act coordinatewise, preserving vanishing of the first two coordinates (Check!).
The sum is all of $\mathbf{F}^5$: given $a = (a_1, \dots, a_5)$, put $u = (a_1, a_2, a_1 + a_2, a_1 - a_2, 2a_1) \in U$ (take $x = a_1$, $y = a_2$); then

$$a - u = \big(0,\, 0,\, a_3 - a_1 - a_2,\, a_4 - a_1 + a_2,\, a_5 - 2a_1\big) \in W,$$

so $a = u + (a - u) \in U + W$. It is direct by 1.46: if $v \in U \cap W$ then $v = (x, y, x + y, x - y, 2x)$ with first two coordinates 0, so $x = y = 0$ and $v = 0$. Hence $\mathbf{F}^5 = U \oplus W$.

Problem 1C.22. Suppose

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbb{F}^5 : x, y \in \mathbb{F}\}.$$

Find three subspaces W_1, W_2, W_3 of $\mathbb{F}^5$, none of which equals $\{0\}$, such that $\mathbb{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

Solution. Take

$$W_1 = \{(0, 0, z, 0, 0) : z \in \mathbb{F}\}, \qquad W_2 = \{(0, 0, 0, z, 0) : z \in \mathbb{F}\}, \qquad W_3 = \{(0, 0, 0, 0, z) : z \in \mathbb{F}\}.$$

Each W_k is a subspace by 1.34, since both operations act only on the single free coordinate (Check!), and none is $\{0\}$, each containing the vector with 1 in that slot. Writing $u(x, y) = (x, y, x+y, x-y, 2x)$, the set U is a subspace too, since

$$u(x, y) + u(x', y') = u(x + x',\, y + y'), \qquad \lambda\, u(x, y) = u(\lambda x, \lambda y),$$

and $0 = u(0, 0)$.
The sum is all of $\mathbb{F}^5$: given $a = (a_1, \dots, a_5)$, put $u = u(a_1, a_2) \in U$; then

$$\begin{aligned} a - u &= (0, 0, a_3 - a_1 - a_2, 0, 0) \\ &\quad + (0, 0, 0, a_4 - a_1 + a_2, 0) + (0, 0, 0, 0, a_5 - 2a_1) \end{aligned}$$

lies in $W_1 + W_2 + W_3$, so $a \in U + W_1 + W_2 + W_3$.
The sum is direct by 1.45: if

$$0 = u(x, y) + (0, 0, c, 0, 0) + (0, 0, 0, d, 0) + (0, 0, 0, 0, e),$$

the first two coordinates give $x = y = 0$, hence $u(x, y) = 0$, and then the third, fourth, and fifth give $c = d = e = 0$. Therefore $\mathbb{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

Problem 1C.23. Prove or give a counterexample: If V_1, V_2, U are subspaces of V such that

$$V = V_1 \oplus U \quad \text{and} \quad V = V_2 \oplus U,$$

then $V_1 = V_2$.
[Hint: When trying to discover whether a conjecture in linear algebra is true or false, it is often useful to start by experimenting in $\mathbb{F}^2$.]

Solution. False. In $V = \mathbb{F}^2$ take

$$U = \{(x, 0) : x \in \mathbb{F}\}, \qquad V_1 = \{(0, y) : y \in \mathbb{F}\}, \qquad V_2 = \{(y, y) : y \in \mathbb{F}\},$$

each a subspace of $\mathbb{F}^2$ by 1.34, its defining condition being linear in the coordinates (Check!). Both sums exhaust $\mathbb{F}^2$, since for $(a, b) \in \mathbb{F}^2$

$$(a, b) = (0, b) + (a, 0) = (b, b) + (a - b, 0),$$

and both are direct by 1.46: a vector of $V_1 \cap U$ has first coordinate 0 (from V_1) and second coordinate 0 (from U), while a vector (a, b) of $V_2 \cap U$ has $b = 0$ from U and $a = b$ from V_2. So $\mathbb{F}^2 = V_1 \oplus U = V_2 \oplus U$, yet $(1, 1) \in V_2 \setminus V_1$, so $V_1 \neq V_2$.

Problem 1C.24. A function $f\colon \mathbb{R} \to \mathbb{R}$ is called even if

$$f(-x) = f(x)$$

for all $x \in \mathbb{R}$. A function $f\colon \mathbb{R} \to \mathbb{R}$ is called odd if

$$f(-x) = -f(x)$$

for all $x \in \mathbb{R}$. Let V_{e} denote the set of real-valued even functions on $\mathbb{R}$ and let V_{o} denote the set of real-valued odd functions on $\mathbb{R}$. Show that $\mathbb{R}^{\mathbb{R}} = V_{\mathrm{e}} \oplus V_{\mathrm{o}}$.

Solution. Each $f \in \mathbb{R}^{\mathbb{R}}$ splits as $f = g + h$ with

$$g(x) = \frac{f(x) + f(-x)}{2} \in V_{\mathrm{e}}, \qquad h(x) = \frac{f(x) - f(-x)}{2} \in V_{\mathrm{o}},$$

since $g(-x) = g(x)$, $h(-x) = -h(x)$, and $g(x) + h(x) = f(x)$ for all x. Hence $\mathbb{R}^{\mathbb{R}} = V_{\mathrm{e}} + V_{\mathrm{o}}$.
Both V_{e} and V_{o} are subspaces by 1.34: the zero function is both even and odd, and since the operations are pointwise,

$$(f + g)(-x) = f(-x) + g(-x), \qquad (\lambda f)(-x) = \lambda f(-x),$$

so each defining sign condition is inherited by sums and scalar multiples (Check!).
The sum is direct by 1.46: if $f \in V_{\mathrm{e}} \cap V_{\mathrm{o}}$ then $f(x) = f(-x) = -f(x)$ for all x, so $2f(x) = 0$ and $f = 0$. Therefore $\mathbb{R}^{\mathbb{R}} = V_{\mathrm{e}} \oplus V_{\mathrm{o}}$.

2 Finite-Dimensional Vector Spaces

Contents

2.1 Exercises 2A

Problem 2A.1. Find a list of four distinct vectors in $\mathbf{F}^3$ whose span equals

$$\{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}.$$

Solution. Take the list

$$(1, 0, -1), \quad (0, 1, -1), \quad (1, -1, 0), \quad (1, 1, -2),$$

whose entries are pairwise distinct (Check!), and write $U = \{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}$. Each listed vector lies in U, its coordinates summing to 0, and U is a subspace of $\mathbf{F}^3$ by 1.34, its defining condition $x + y + z = 0$ being linear in the coordinates (Check!); so by 2.6, which makes the span the smallest subspace containing the list, the span lies in U. Conversely, if $(x, y, z) \in U$ then $z = -x - y$, so

$$(x, y, z) = x(1, 0, -1) + y(0, 1, -1) + 0\,(1, -1, 0) + 0\,(1, 1, -2)$$

lies in the span. Hence the span equals U.

Problem 2A.2. Prove or give a counterexample: If v_1, v_2, v_3, v_4 spans V, then the list

$$v_1 - v_2, \quad v_2 - v_3, \quad v_3 - v_4, \quad v_4$$

also spans V.

Solution. True. Put $W = \operatorname{span}(v_1 - v_2,\ v_2 - v_3,\ v_3 - v_4,\ v_4)$, a subspace of V by 2.6. Telescoping from the last vector backwards,

$$\begin{aligned}
v_4 &= v_4, \\
v_3 &= (v_3 - v_4) + v_4, \\
v_2 &= (v_2 - v_3) + (v_3 - v_4) + v_4, \\
v_1 &= (v_1 - v_2) + (v_2 - v_3) + (v_3 - v_4) + v_4,
\end{aligned}$$

so $v_1, v_2, v_3, v_4 \in W$. By 2.6, $\operatorname{span}(v_1, v_2, v_3, v_4)$ is the smallest subspace containing them, whence $V = \operatorname{span}(v_1, v_2, v_3, v_4) \subseteq W \subseteq V$ and $W = V$.

Problem 2A.3. Suppose $v_1, \ldots, v_m$ is a list of vectors in V. For $k \in \{1, \ldots, m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that $\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(w_1, \ldots, w_m)$.

Solution. Each list lies in the span of the other, so by the minimality in 2.6 the two spans contain each other. Writing $A = \operatorname{span}(v_1, \ldots, v_m)$ and $B = \operatorname{span}(w_1, \ldots, w_m)$, both subspaces of V by 2.6:

$$w_k = 1\,v_1 + \cdots + 1\,v_k + 0\,v_{k+1} + \cdots + 0\,v_m \in A,$$

so $B \subseteq A$; and $v_1 = w_1 \in B$ while for $k \geq 2$

$$v_k = (v_1 + \cdots + v_k) - (v_1 + \cdots + v_{k-1}) = w_k - w_{k-1} \in B,$$

so $A \subseteq B$. Hence $\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(w_1, \ldots, w_m)$.

Problem 2A.4. (a) Show that a list of length one in a vector space is linearly independent if and only if the vector in the list is not 0.
(b) Show that a list of length two in a vector space is linearly independent if and only if neither of the two vectors in the list is a scalar multiple of the other.

Solution. (a) The list v of length one is linearly independent exactly when $v \neq 0$.
If $v \neq 0$ and $av = 0$ with $a \neq 0$, then a is invertible and

$$v = \big(a^{-1}a\big)v = a^{-1}(av) = a^{-1}\,0 = 0$$

by 1.31, a contradiction; so $a = 0$ and the list is linearly independent. If instead $v = 0$, then $1 \cdot v = 0$ with $1 \neq 0$, so the list is linearly dependent (2.17).
(b) The list v, w is linearly dependent exactly when one of v, w is a scalar multiple of the other; the assertion is the contrapositive.
Suppose v, w is linearly dependent, so $av + bw = 0$ with a, b not both 0 (2.17).
(i) $a \neq 0$: multiplying by a^{-1} gives $v = -\big(a^{-1}b\big)w$.
(ii) $a = 0$: then $b \neq 0$, so $w = b^{-1}(bw) = 0 = 0\,v$ by 1.30.
Conversely, if $v = \lambda w$ then $1 \cdot v + (-\lambda)w = 0$, and if $w = \mu v$ then $(-\mu)v + 1 \cdot w = 0$; the coefficient $1 \neq 0$ makes each combination nontrivial, so the list is linearly dependent (2.17).

Problem 2A.5. Find a number t such that

$$(3, 1, 4), \quad (2, -3, 5), \quad (5, 9, t)$$

is not linearly independent in $\mathbf{R}^3$.

Solution. Take $t = 2$, for which

$$(-3)(3, 1, 4) + 2(2, -3, 5) + (5, 9, 2) = (-9 + 4 + 5,\ -3 - 6 + 9,\ -12 + 10 + 2) = (0, 0, 0).$$

The scalars $-3, 2, 1$ are not all 0, so the list is linearly dependent (2.17), i.e. not linearly independent.
Method (2): solving $a(3,1,4)+b(2,-3,5)+(5,9,t) = 0$ for the first two coordinates gives $a = 3b-9$ and $3(3b-9)+2b+5 = 11b-22 = 0$, so $b = 2$ and $a = -3$; the third coordinate then reads $-12+10+t = 0$, i.e. $t = 2$.

Problem 2A.6. Show that the list $(2, 3, 1), (1, -1, 2), (7, 3, c)$ is linearly dependent in $\mathbf{F}^3$ if and only if $c = 8$.

Solution. If $c = 8$ the list is linearly dependent, because

$$2(2, 3, 1) + 3(1, -1, 2) + (-1)(7, 3, 8) = (4 + 3 - 7,\ 6 - 3 - 3,\ 2 + 6 - 8) = (0, 0, 0)$$

with scalars not all 0 (2.17).
Conversely, suppose $a(2, 3, 1) + b(1, -1, 2) + d(7, 3, c) = (0, 0, 0)$ with a, b, d not all 0 (2.17). Here $(2, 3, 1), (1, -1, 2)$ is linearly independent by Exercise 2A.4(b): in $(1, -1, 2) = \lambda(2, 3, 1)$ the third coordinates force $\lambda = 2$ while the first force $2\lambda = 1$, and symmetrically for the other multiple. Hence $d \neq 0$, since $d = 0$ would force $a = b = 0$ as well. Dividing by d and putting $\alpha = -a/d$, $\beta = -b/d$ gives $(7, 3, c) = \alpha(2, 3, 1) + \beta(1, -1, 2)$, that is,

$$2\alpha + \beta = 7, \qquad 3\alpha - \beta = 3, \qquad \alpha + 2\beta = c.$$

Adding the first two gives $\alpha = 2$, then $\beta = 3$, and so $c = 2 + 6 = 8$.

Problem 2A.7. (a) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{R}$, then the list $1+i,\ 1-i$ is linearly independent.
(b) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{C}$, then the list $1+i,\ 1-i$ is linearly dependent.

Solution. (a) Only real scalars are allowed, so suppose $a, b \in \mathbf{R}$ satisfy

$$a(1+i) + b(1-i) = (a+b) + (a-b)i = 0.$$

A complex number vanishes exactly when its real and imaginary parts do, and $a \pm b$ are real, so $a+b=0$ and $a-b=0$; adding and subtracting gives $a=b=0$. The list is linearly independent (2.15).
(b) Now $-i$ is an allowed scalar, and

$$(-i)(1+i) + (-1)(1-i) = (1-i) - (1-i) = 0$$

is a nontrivial vanishing combination, so the list is linearly dependent (2.17).

Problem 2A.8. Suppose v_1, v_2, v_3, v_4 is linearly independent in V. Prove that the list

$$v_1 - v_2, \quad v_2 - v_3, \quad v_3 - v_4, \quad v_4$$

is also linearly independent.

Solution. Suppose $a_1, a_2, a_3, a_4 \in \mathbb{F}$ satisfy

$$\begin{aligned} 0 &= a_1(v_1 - v_2) + a_2(v_2 - v_3) + a_3(v_3 - v_4) + a_4 v_4 \\ &= a_1 v_1 + (a_2 - a_1)v_2 + (a_3 - a_2)v_3 + (a_4 - a_3)v_4. \end{aligned}$$

Linear independence of v_1, v_2, v_3, v_4 forces every coefficient to vanish (2.15):

$$a_1 = 0, \qquad a_2 - a_1 = 0, \qquad a_3 - a_2 = 0, \qquad a_4 - a_3 = 0,$$

and reading these in order gives $a_1 = a_2 = a_3 = a_4 = 0$. Hence the list is linearly independent.

Problem 2A.9. Prove or give a counterexample: If $v_1, v_2, \ldots, v_m$ is a linearly independent list of vectors in V, then

$$5v_1 - 4v_2, \quad v_2, \quad v_3, \quad \ldots, \quad v_m$$

is linearly independent.

Solution. True. Suppose $a_1, \ldots, a_m \in \mathbb{F}$ satisfy

$$\begin{aligned} 0 &= a_1(5v_1 - 4v_2) + a_2 v_2 + a_3 v_3 + \cdots + a_m v_m \\ &= (5a_1)v_1 + (a_2 - 4a_1)v_2 + a_3 v_3 + \cdots + a_m v_m. \end{aligned}$$

Linear independence of $v_1, \ldots, v_m$ makes every coefficient 0 (2.15):

$$5a_1 = 0, \qquad a_2 - 4a_1 = 0, \qquad a_3 = \cdots = a_m = 0.$$

Since $\mathbb{F}$ is $\mathbb{R}$ or $\mathbb{C}$, we have $5 \neq 0$, so $a_1 = 0$ and then $a_2 = 4a_1 = 0$. Hence all the scalars are 0 and the list is linearly independent.

Problem 2A.10. Prove or give a counterexample: If $v_1, v_2, \ldots, v_m$ is a linearly independent list of vectors in V and $\lambda \in \mathbb{F}$ with $\lambda \neq 0$, then $\lambda v_1, \lambda v_2, \ldots, \lambda v_m$ is linearly independent.

Solution. True. Suppose $a_1, \dots, a_m \in \mathbb{F}$ satisfy

$$0 = a_1(\lambda v_1) + \cdots + a_m(\lambda v_m) = (a_1\lambda)v_1 + \cdots + (a_m\lambda)v_m,$$

the second equality by commutativity and associativity of scalar multiplication. Linear independence of $v_1, \dots, v_m$ gives $a_k\lambda = 0$ for every k (2.15), and $\lambda \neq 0$ is invertible in $\mathbb{F}$, so $a_k = \lambda^{-1}(a_k\lambda) = 0$ for every k. Hence $\lambda v_1, \dots, \lambda v_m$ is linearly independent.

Problem 2A.11. Prove or give a counterexample: If $v_1, \dots, v_m$ and $w_1, \dots, w_m$ are linearly independent lists of vectors in V, then the list $v_1 + w_1, \dots, v_m + w_m$ is linearly independent.

Solution. False. In $V = \mathbb{R}^2$ with $m = 2$, take

$$v_1 = (1, 0), \quad v_2 = (0, 1), \qquad w_1 = (-1, 0), \quad w_2 = (0, -1).$$

Each of v_1, v_2 and w_1, w_2 is linearly independent, since $a_1(1,0) + a_2(0,1) = (a_1, a_2)$ and $b_1(-1,0) + b_2(0,-1) = (-b_1, -b_2)$ vanish only for zero scalars. But $v_1 + w_1 = v_2 + w_2 = (0,0)$, and a list containing 0 is linearly dependent (2.18).

Problem 2A.12. Suppose $v_1, \dots, v_m$ is linearly independent in V and $w \in V$. Prove that if $v_1 + w, \dots, v_m + w$ is linearly dependent, then $w \in \operatorname{span}(v_1, \dots, v_m)$.

Solution. By 2.17 there are $a_1, \dots, a_m \in \mathbb{F}$, not all 0, with

$$0 = \sum_{k=1}^{m} a_k(v_k + w) = a_1v_1 + \cdots + a_mv_m + sw, \qquad s = a_1 + \cdots + a_m.$$

Here $s \neq 0$: otherwise $a_1v_1 + \cdots + a_mv_m = 0$ with the a_k not all 0, contradicting linear independence of $v_1, \dots, v_m$ (2.15). Dividing by s,

$$w = -(a_1/s)v_1 - \cdots - (a_m/s)v_m \in \operatorname{span}(v_1, \dots, v_m).$$

Problem 2A.13. Suppose $v_1, \dots, v_m$ is linearly independent in V and $w \in V$. Show that

$$v_1, \dots, v_m, w \text{ is linearly independent} \iff w \notin \operatorname{span}(v_1, \dots, v_m).$$

Solution. If $w \in \operatorname{span}(v_1, \dots, v_m)$, say $w = c_1v_1 + \cdots + c_mv_m$, then

$$c_1v_1 + \cdots + c_mv_m + (-1)w = 0$$

has last scalar $-1 \neq 0$, so $v_1, \dots, v_m, w$ is linearly dependent (2.17). That is the contrapositive of one implication.

Conversely, suppose $w \notin \operatorname{span}(v_1, \dots, v_m)$ and $a_1v_1 + \cdots + a_mv_m + bw = 0$. Were $b \neq 0$, inverting b would give

$$w = -(a_1/b)v_1 - \cdots - (a_m/b)v_m \in \operatorname{span}(v_1, \dots, v_m),$$

contrary to hypothesis. So $b = 0$, leaving $a_1v_1 + \cdots + a_mv_m = 0$, whence $a_1 = \cdots = a_m = 0$ by linear independence of $v_1, \dots, v_m$ (2.15). All the scalars vanish, so $v_1, \dots, v_m, w$ is linearly independent.

Problem 2A.14. Suppose $v_1, \dots, v_m$ is a list of vectors in V. For $k \in \{1, \dots, m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that the list $v_1, \dots, v_m$ is linearly independent if and only if the list $w_1, \dots, w_m$ is linearly independent.

Solution. Each list is recovered from the other: $v_1 = w_1$ and $v_k = w_k - w_{k-1}$ for $2 \le k \le m$, by cancellation in the partial sums.
(i) Suppose $v_1, \dots, v_m$ is linearly independent and $a_1 w_1 + \cdots + a_m w_m = 0$. Since v_j occurs in w_k exactly when $k \ge j$,

$$0 = \sum_{k=1}^{m} a_k w_k = \sum_{j=1}^{m} \Big(\sum_{k=j}^{m} a_k \Big) v_j,$$

so $\sum_{k=j}^{m} a_k = 0$ for each j (2.15); subtracting the equation for $j+1$ from that for j gives $a_j = 0$ for $j < m$, and $j = m$ gives $a_m = 0$. Hence $w_1, \dots, w_m$ is linearly independent.
(ii) Suppose $w_1, \dots, w_m$ is linearly independent and $b_1 v_1 + \cdots + b_m v_m = 0$. Substituting $v_1 = w_1$ and $v_k = w_k - w_{k-1}$,

$$0 = \sum_{k=1}^{m} b_k w_k - \sum_{k=1}^{m-1} b_{k+1} w_k = \sum_{k=1}^{m-1} (b_k - b_{k+1}) w_k + b_m w_m,$$

so $b_1 = b_2 = \cdots = b_m$ and $b_m = 0$ (2.15); hence all $b_k = 0$ and $v_1, \dots, v_m$ is linearly independent.

Problem 2A.15. Explain why there does not exist a list of six polynomials that is linearly independent in $\mathcal{P}_4(\mathbf{F})$.

Solution. Every linearly independent list in $\mathcal{P}_4(\mathbf{F})$ has length at most 5, so no list of six can be linearly independent.
Indeed $1, z, z^2, z^3, z^4$ spans $\mathcal{P}_4(\mathbf{F})$, since each $p \in \mathcal{P}_4(\mathbf{F})$ has $\deg p \le 4$ and so is written as $p(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4$ by 2.11 (padding with zero coefficients; $p = 0$ takes all $a_k = 0$). Thus $\mathcal{P}_4(\mathbf{F})$ is finite-dimensional (2.9) with a spanning list of length 5, and 2.22 bounds every linearly independent list by that length.

Problem 2A.16. Explain why no list of four polynomials spans $\mathcal{P}_4(\mathbf{F})$.

Solution. The list $1, z, z^2, z^3, z^4$ is linearly independent in $\mathcal{P}_4(\mathbf{F})$ and has length 5, so by 2.22 no spanning list of length 4 can exist.
Linear independence is 2.16(b): if $a_0 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4 = 0$ for every $z \in \mathbf{F}$, then all $a_k = 0$. (That condition involves only the listed vectors and $\mathbf{F}$, so it reads the same in the subspace $\mathcal{P}_4(\mathbf{F})$ as in $\mathcal{P}(\mathbf{F})$.) A spanning list q_1, q_2, q_3, q_4 would make $\mathcal{P}_4(\mathbf{F})$ finite-dimensional (2.9) with $5 \le 4$ by 2.22, which is false.

Problem 2A.17. Prove that V is infinite-dimensional if and only if there is a sequence $v_1, v_2, \dots$ of vectors in V such that $v_1, \dots, v_m$ is linearly independent for every positive integer m.

Solution. ($\Leftarrow$) Given such a sequence, a spanning list $w_1, \dots, w_n$ of V (2.9) would force $n + 1 \le n$ by 2.22 applied to the linearly independent list $v_1, \dots, v_{n+1}$. So V is not finite-dimensional, i.e. infinite-dimensional (2.13).
($\Rightarrow$) Suppose V is infinite-dimensional and choose v_m recursively.
(i) $V \ne \{0\}$, since otherwise the empty list spans V; pick $v_1 \ne 0$, linearly independent by 2.16(c).
(ii) Given $v_1, \dots, v_m$ linearly independent, no list spans V, so $\operatorname{span}(v_1, \dots, v_m) \ne V$ and we may choose

$$v_{m+1} \in V \setminus \operatorname{span}(v_1, \dots, v_m).$$

By Exercise 2A.13 the list $v_1, \dots, v_{m+1}$ is linearly independent.
The recursion never halts, and every initial segment is linearly independent.

Problem 2A.18. Prove that $\mathbf{F}^\infty$ is infinite-dimensional.

Solution. Take $e_k \in \mathbf{F}^\infty$ to be the sequence with 1 in coordinate k and 0 elsewhere. For every m,

$$a_1e_1 + \cdots + a_me_m = (a_1, \ldots, a_m, 0, 0, \ldots),$$

which is the zero sequence only if $a_1 = \cdots = a_m = 0$; so $e_1, \ldots, e_m$ is linearly independent. By Exercise 2A.17, $\mathbf{F}^\infty$ is infinite-dimensional.

Problem 2A.19. Prove that the real vector space of all continuous real-valued functions on the interval $[0, 1]$ is infinite-dimensional.

Solution. Take $f_k \in C[0, 1]$ to be $f_k(x) = x^k$ for $k = 0, 1, 2, \ldots$, each continuous on $[0, 1]$.
For every m the list $f_0, \ldots, f_m$ is linearly independent: if $a_0f_0 + \cdots + a_mf_m = 0$ in $C[0, 1]$, then the polynomial $p(z) = a_0 + a_1z + \cdots + a_mz^m$ vanishes at every point of $[0, 1]$, hence has infinitely many zeros in $\mathbf{R}$, so $p = 0$ by 4.8 (a nonzero p of degree d has at most d zeros for $d \geq 1$, and none for $d = 0$). Uniqueness of polynomial coefficients then gives $a_0 = \cdots = a_m = 0$.
Applying Exercise 2A.17 to the sequence $f_0, f_1, f_2, \ldots$ shows $C[0, 1]$ is infinite-dimensional.

Problem 2A.20. Suppose $p_0, p_1, \ldots, p_m$ are polynomials in $\mathcal{P}_m(\mathbf{F})$ such that $p_k(2) = 0$ for each $k \in \{0, \ldots, m\}$. Prove that $p_0, p_1, \ldots, p_m$ is not linearly independent in $\mathcal{P}_m(\mathbf{F})$.

Solution. The $m + 1$ polynomials all lie in

$$U = \{p \in \mathcal{P}_m(\mathbf{F}) : p(2) = 0\},$$

which is spanned by a list of length m; so 2.22 forbids their linear independence.
U is a subspace of $\mathcal{P}_m(\mathbf{F})$ by 1.34, since $0 \in U$ and $(p + \lambda q)(2) = p(2) + \lambda q(2) = 0$ for $p, q \in U$. For $j \in \{1, \ldots, m\}$ put $q_j(z) = z^j - 2^j$; each has degree $j \leq m$ and $q_j(2) = 0$, so $q_j \in U$ and $\operatorname{span}(q_1, \ldots, q_m) \subseteq U$. Conversely, writing $p \in U$ as $p(z) = a_0 + a_1z + \cdots + a_mz^m$ (2.11, 2.12),

$$p(z) = p(z) - p(2) = a_1q_1(z) + a_2q_2(z) + \cdots + a_mq_m(z),$$

so $U = \operatorname{span}(q_1, \ldots, q_m)$ and U is finite-dimensional (2.9). (For $m = 0$ both sides are $\{0\}$.)
Linear independence depends only on the vectors and $\mathbf{F}$, so were $p_0, \ldots, p_m$ linearly independent in $\mathcal{P}_m(\mathbf{F})$ it would be a linearly independent list of length $m + 1$ in U, giving $m + 1 \leq m$ by 2.22.

2.2 Exercises 2B

Problem 2B.1. Find all vector spaces that have exactly one basis.

Solution. Only the zero vector space $\{0\}$, whose unique basis is the empty list.
The empty list is linearly independent (2.15) with span $\{0\}$ (2.4), hence a basis of $\{0\}$ by 2.26; every nonempty list in $\{0\}$ reads $0, \ldots, 0$ and is linearly dependent (2.18). So $\{0\}$ has exactly one basis.
Conversely, let $v_1, \ldots, v_n$ be the unique basis of V and suppose $n \geq 1$. Then $v_1 \neq 0$ by linear independence, and $2v_1, v_2, \ldots, v_n$ is again a basis: it spans, since $v_1 = \frac{1}{2}(2v_1)$ puts $V = \operatorname{span}(v_1, \ldots, v_n)$ inside its span ($\frac{1}{2} \in \mathbf{F}$ as $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$); and $a_1(2v_1) + a_2v_2 + \cdots + a_nv_n = 0$ gives $2a_1 = a_2 = \cdots = a_n = 0$, so all $a_k = 0$. But $2v_1 \neq v_1$ since $2v_1 - v_1 = v_1 \neq 0$, so this is a second basis, a contradiction. Hence $n = 0$ and $V = \operatorname{span}() = \{0\}$.

Problem 2B.2. Verify all assertions in Example 2.27.

Solution. Throughout, a list is a basis exactly when it is linearly independent and spans (2.26).
(a) $a_1e_1 + \cdots + a_ne_n = (a_1, \ldots, a_n)$, so every $(x_1, \ldots, x_n) \in \mathbf{F}^n$ equals $x_1e_1 + \cdots + x_ne_n$ and the combination vanishes only for $a_1 = \cdots = a_n = 0$. Hence $e_1, \ldots, e_n$ is a basis of $\mathbf{F}^n$.
(b) $a(1, 2) + b(3, 5) = (x, y)$ reads $a + 3b = x$, $2a + 5b = y$, whose unique solution is $b = 2x - y$, $a = -5x + 3y$; and

$$(-5x + 3y)(1, 2) + (2x - y)(3, 5) = (x, y). \quad \text{(Check!)}$$

So representations exist and are unique, making $(1, 2), (3, 5)$ a basis of $\mathbf{F}^2$ by 2.28, of length two like the standard basis in (a).
(c) $a(1, 2, -4) + b(7, -5, 6) = (0, 0, 0)$ gives $a + 7b = 0$ and $2a - 5b = 0$, hence $-19b = 0$ and $a = b = 0$: linearly independent. It does not span, since $a(1, 2, -4) + b(7, -5, 6) = (0, 0, 1)$ forces $a = b = 0$ from the first two coordinates and then $0 \neq 1$ in the third. So it is not a basis of $\mathbf{F}^3$.
(d) It spans $\mathbf{F}^2$, since the sublist $(1, 2), (3, 5)$ already does by (b) and enlarging a list only enlarges its span. It is linearly dependent: the formulas of (b) at $(x, y) = (4, 13)$ give

$$19(1, 2) - 5(3, 5) - 1 \cdot (4, 13) = (0, 0),$$

a nontrivial vanishing combination (2.17). Hence not a basis.
(e) $a(1, 1, 0) + b(0, 0, 1) = (a, a, b)$, which lies in $U = \{(x, x, y)\}$, and conversely $(x, x, y) = x(1, 1, 0) + y(0, 0, 1)$; so $U = \operatorname{span}\big((1, 1, 0), (0, 0, 1)\big)$ is a subspace (2.6) that the list spans. It vanishes only for $a = b = 0$, so the list is a basis of U.
(f) $a(1, -1, 0) + b(1, 0, -1) = (a + b, -a, -b)$, whose coordinates sum to 0, so the span lies in $U = \{x + y + z = 0\}$; conversely $(x, y, z) \in U$ has $x = -y - z$ and

$$(-y)(1, -1, 0) + (-z)(1, 0, -1) = (-y - z,\ y,\ z) = (x, y, z).$$

Hence U is that span, a subspace by 2.6; and $(a + b, -a, -b) = 0$ forces $a = b = 0$. So the list is a basis of U.
(g) By definition every $p \in \mathcal{P}_m(\mathbf{F})$ is $p(z) = a_0 + a_1z + \cdots + a_mz^m$, i.e. $1, z, \ldots, z^m$ spans $\mathcal{P}_m(\mathbf{F})$; it is linearly independent by 2.16(b). Hence it is a basis.
Finally, $(7, 5), (-4, 9)$ is a basis of $\mathbf{F}^2$: the system $7a - 4b = x$, $5a + 9b = y$ has determinant $83 \neq 0$ in $\mathbf{F}$, so it has the unique solution

$$a = \frac{9x + 4y}{83}, \qquad b = \frac{7y - 5x}{83} \quad \text{(Check!)},$$

and 2.28 applies. That $(1, 2), (3, 5)$ is a basis was shown in (b).

Problem 2B.3. (a) Let U be the subspace of $\mathbf{R}^5$ defined by

$$U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{R}^5 : x_1 = 3x_2 \text{ and } x_3 = 7x_4\}.$$

Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathbf{R}^5$.
(c) Find a subspace W of $\mathbf{R}^5$ such that $\mathbf{R}^5 = U \oplus W$.

Solution. (a) A basis of U is

$$(3, 1, 0, 0, 0), \quad (0, 0, 7, 1, 0), \quad (0, 0, 0, 0, 1),$$

each vector lying in U (Check!). It spans: the constraints $x_1 = 3x_2$, $x_3 = 7x_4$ leave $x_2 = t$, $x_4 = s$, $x_5 = r$ free, and

$$(3t, t, 7s, s, r) = t(3, 1, 0, 0, 0) + s(0, 0, 7, 1, 0) + r(0, 0, 0, 0, 1).$$

It is linearly independent, since that combination equals $(3a, a, 7b, b, c)$, whose second, fourth, and fifth coordinates give $a = b = c = 0$. So it is a basis by 2.26.

(b) Adjoin $(1,0,0,0,0)$ and $(0,0,1,0,0)$, giving the list

$$(3,1,0,0,0),\ (0,0,7,1,0),\ (0,0,0,0,1),\ (1,0,0,0,0),\ (0,0,1,0,0).$$

The combination $a(3,1,0,0,0) + b(0,0,7,1,0) + c(0,0,0,0,1) + d(1,0,0,0,0) + e(0,0,1,0,0)$ equals $(3a+d,\ a,\ 7b+e,\ b,\ c)$. It spans $\mathbf{R}^5$: take $a = x_2$, $b = x_4$, $c = x_5$, $d = x_1 - 3x_2$, $e = x_3 - 7x_4$. It is linearly independent: if that vector is 0, coordinates two, four, five give $a = b = c = 0$, and then coordinates one and three give $d = e = 0$. So this length-five list is a basis of $\mathbf{R}^5$ extending (a).
(c) Take

$$\begin{aligned} W &= \operatorname{span}\big((1,0,0,0,0),(0,0,1,0,0)\big) \\ &= \{(x_1,0,x_3,0,0) : x_1, x_3 \in \mathbf{R}\}, \end{aligned}$$

a subspace by 2.6. By (b) every vector of $\mathbf{R}^5$ splits as a combination of the first three listed vectors, which lies in U, plus one of the last two, which lies in W; so $U + W = \mathbf{R}^5$. And $(x_1,0,x_3,0,0) \in U$ forces $x_1 = 3 \cdot 0 = 0$ and $x_3 = 7 \cdot 0 = 0$, so $U \cap W = \{0\}$. Hence $\mathbf{R}^5 = U \oplus W$ by 1.46.

Problem 2B.4. (a) Let U be the subspace of $\mathbf{C}^5$ defined by

$$U = \{(z_1,z_2,z_3,z_4,z_5) \in \mathbf{C}^5 : 6z_1 = z_2 \text{ and } z_3 + 2z_4 + 3z_5 = 0\}.$$

Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathbf{C}^5$.
(c) Find a subspace W of $\mathbf{C}^5$ such that $\mathbf{C}^5 = U \oplus W$.

Solution. (a) A basis of U is

$$(1,6,0,0,0), \quad (0,0,-2,1,0), \quad (0,0,-3,0,1),$$

each vector lying in U (Check!). It spans: the constraints $z_2 = 6z_1$, $z_3 = -2z_4 - 3z_5$ leave $z_1 = t$, $z_4 = s$, $z_5 = r$ free, and

$$(t, 6t, -2s-3r, s, r) = t(1,6,0,0,0) + s(0,0,-2,1,0) + r(0,0,-3,0,1).$$

It is linearly independent, since that combination equals $(a, 6a, -2b-3c, b, c)$, whose first, fourth, and fifth coordinates give $a = b = c = 0$. So it is a basis by 2.26.
(b) Adjoin $(0,1,0,0,0)$ and $(0,0,1,0,0)$, giving

$$(1,6,0,0,0),\ (0,0,-2,1,0),\ (0,0,-3,0,1),\ (0,1,0,0,0),\ (0,0,1,0,0).$$

The combination $a(1,6,0,0,0) + b(0,0,-2,1,0) + c(0,0,-3,0,1) + d(0,1,0,0,0) + e(0,0,1,0,0)$ equals $(a,\ 6a+d,\ -2b-3c+e,\ b,\ c)$. It spans $\mathbf{C}^5$: take $a = z_1$, $b = z_4$, $c = z_5$, $d = z_2 - 6z_1$, $e = z_3 + 2z_4 + 3z_5$. It is linearly independent: if that vector is 0, coordinates one, four, five give $a = b = c = 0$, and then coordinates two and three give $d = e = 0$. So it is a basis of $\mathbf{C}^5$ extending (a).
(c) Take

$$\begin{aligned} W &= \operatorname{span}\big((0,1,0,0,0),(0,0,1,0,0)\big) \\ &= \{(0,z_2,z_3,0,0) : z_2, z_3 \in \mathbf{C}\}, \end{aligned}$$

a subspace by 2.6. By (b) every vector of $\mathbf{C}^5$ splits as a combination of the first three listed vectors, which lies in U, plus one of the last two, which lies in W; so $U + W = \mathbf{C}^5$. And $(0,z_2,z_3,0,0) \in U$ forces $z_2 = 6 \cdot 0 = 0$ and $z_3 = 0$, so $U \cap W = \{0\}$. Hence $\mathbf{C}^5 = U \oplus W$ by 1.46.

Problem 2B.5. Suppose V is finite-dimensional and U, W are subspaces of V such that $V = U + W$. Prove that there exists a basis of V consisting of vectors in $U \cup W$.

Solution. Reduce the concatenated bases of U and W to a basis of V (2.30).
In detail: U and W are finite-dimensional (2.25), so they have bases $u_1, \ldots, u_m$ and $w_1, \ldots, w_n$ (2.31). Each $v \in V$ is $v = u + w$ with $u \in U$, $w \in W$ since $V = U + W$, and expanding u and w in those bases gives

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_n,$$

so $u_1, \ldots, u_m, w_1, \ldots, w_n$ spans V. By 2.30 some sublist of it is a basis of V, and every vector of that sublist lies in $U \cup W$.

Problem 2B.6. Prove or give a counterexample: If p_0, p_1, p_2, p_3 is a list in $\mathcal{P}_3(\mathbf{F})$ such that none of the polynomials p_0, p_1, p_2, p_3 has degree 2, then p_0, p_1, p_2, p_3 is not a basis of $\mathcal{P}_3(\mathbf{F})$.

Solution. False: take

$$p_0(z) = 1, \qquad p_1(z) = z, \qquad p_2(z) = z^2 + z^3, \qquad p_3(z) = z^3,$$

of degrees $0, 1, 3, 3$, none equal to 2.
This list spans, since

$$a + bz + cz^2 + dz^3 = a\,p_0 + b\,p_1 + c\,p_2 + (d - c)\,p_3,$$

and $1, z, z^2, z^3$ spans $\mathcal{P}_3(\mathbf{F})$ (Example 2.27(g)). It is linearly independent: $a\,p_0 + b\,p_1 + c\,p_2 + d\,p_3 = 0$ reads $a + bz + cz^2 + (c + d)z^3 = 0$ for all z, so $a = b = c = 0$ and $c + d = 0$ by 2.16(b), whence $d = 0$. So it is a basis by 2.26.

Problem 2B.7. Suppose v_1, v_2, v_3, v_4 is a basis of V. Prove that

$$v_1 + v_2,\ v_2 + v_3,\ v_3 + v_4,\ v_4$$

is also a basis of V.

Solution. By 2.26 it suffices that the list be linearly independent and span V.
Linear independence: regrouping $a(v_1 + v_2) + b(v_2 + v_3) + c(v_3 + v_4) + dv_4 = 0$ gives

$$av_1 + (a + b)v_2 + (b + c)v_3 + (c + d)v_4 = 0,$$

so $a = a + b = b + c = c + d = 0$ by linear independence of the basis v_1, v_2, v_3, v_4; reading in order, $a = b = c = d = 0$.
Spanning: writing S for the span of the new list, which is a subspace (2.6),

$$v_4 \in S, \quad v_3 = (v_3 + v_4) - v_4 \in S, \quad v_2 = (v_2 + v_3) - v_3 \in S,$$

and then $v_1 = (v_1 + v_2) - v_2 \in S$. Hence $V = \operatorname{span}(v_1, v_2, v_3, v_4) \subseteq S$, so $S = V$.

Problem 2B.8. Prove or give a counterexample: If v_1, v_2, v_3, v_4 is a basis of V and U is a subspace of V such that $v_1, v_2 \in U$ and $v_3 \notin U$ and $v_4 \notin U$, then v_1, v_2 is a basis of U.

Solution. False: in $V = \mathbf{F}^4$ take v_1, v_2, v_3, v_4 to be the standard basis e_1, e_2, e_3, e_4 and

$$\begin{aligned} U &= \operatorname{span}(e_1,\ e_2,\ e_3 + e_4) \\ &= \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_3 = x_4\}, \end{aligned}$$

the two descriptions agreeing because $e_1, e_2, e_3 + e_4$ all satisfy $x_3 = x_4$, a subspace condition, while $x_3 = x_4$ gives $(x_1, x_2, x_3, x_4) = x_1 e_1 + x_2 e_2 + x_3(e_3 + e_4)$.
The hypotheses hold: $e_1, e_2 \in U$ (third and fourth coordinates both 0), while $e_3, e_4 \notin U$ (their third and fourth coordinates differ). But $e_3 + e_4 \in U$ has third coordinate 1, whereas every vector in $\operatorname{span}(e_1, e_2)$ has third coordinate 0; so v_1, v_2 does not span U and is not a basis of U (2.26).

Problem 2B.9. Suppose $v_1, \ldots, v_m$ is a list of vectors in V. For $k \in \{1, \ldots, m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that $v_1, \ldots, v_m$ is a basis of V if and only if $w_1, \ldots, w_m$ is a basis of V.

Solution. With $w_0 = 0$ one has $v_k = w_k - w_{k-1}$ for every k, so each list lies in the span of the other and

$$\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(w_1, \ldots, w_m);$$

in particular one list spans V exactly when the other does. Expanding each side in the other list (with $b_{m+1} = 0$),

$$\sum_{k=1}^{m} a_k w_k = \sum_{j=1}^{m} \Big(\sum_{k=j}^{m} a_k\Big) v_j,$$
$$\sum_{k=1}^{m} b_k v_k = \sum_{k=1}^{m} (b_k - b_{k+1}) w_k.$$

(i) If $v_1, \ldots, v_m$ is linearly independent and $\sum_k a_k w_k = 0$, the first identity forces $\sum_{k=j}^{m} a_k = 0$ for every j; subtracting the equation for $j+1$ from that for j gives $a_j = 0$.
(ii) If $w_1, \ldots, w_m$ is linearly independent and $\sum_k b_k v_k = 0$, the second identity forces $b_k = b_{k+1}$ for every k, and $b_{m+1} = 0$ then gives $b_1 = \cdots = b_m = 0$.
So the lists are linearly independent together, and by 2.26 one is a basis of V exactly when the other is.

Problem 2B.10. Suppose U and W are subspaces of V such that $V = U \oplus W$. Suppose also that $u_1, \ldots, u_m$ is a basis of U and $w_1, \ldots, w_n$ is a basis of W. Prove that

$$u_1, \ldots, u_m, w_1, \ldots, w_n$$

is a basis of V.

Solution. By 2.26 it suffices to check spanning and linear independence of the concatenated list.
Spanning. Each $v \in V$ is $v = u + w$ with $u \in U$, $w \in W$, since $V = U + W$; expanding u and w in their bases gives

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_n.$$

Independence. If that sum is 0, then $u + w = 0$ with $u = \sum_j a_j u_j \in U$ and $w = \sum_k b_k w_k \in W$, so $u = w = 0$ by 1.45 (the sum $U + W$ is direct); independence of each of the two bases makes every a_j and every b_k zero.

Problem 2B.11. Suppose V is a real vector space. Show that if $v_1, \ldots, v_n$ is a basis of V (as a real vector space), then $v_1, \ldots, v_n$ is also a basis of the complexification $V_{\mathbf{C}}$ (as a complex vector space). See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$.

Solution. Everything follows from the identity $(*)$: for $a_k, b_k \in \mathbf{R}$, with each v_k read as $v_k + i0$,

$$\sum_{k=1}^{n} (a_k + b_k i)(v_k + i0) = \Big(\sum_{k=1}^{n} a_k v_k\Big) + i\Big(\sum_{k=1}^{n} b_k v_k\Big),$$

since $(a + bi)(v + i0) = (av) + i(bv)$ and addition in $V_{\mathbf{C}}$ is componentwise (Exercise 8 in Section 1B). As every complex scalar is $a + bi$ with $a, b \in \mathbf{R}$, $(*)$ covers all complex linear combinations of the list.
Spanning. Given $u + iv \in V_{\mathbf{C}}$, expand $u = \sum_k a_k v_k$ and $v = \sum_k b_k v_k$ over $\mathbf{R}$; then $(*)$ gives $\sum_k (a_k + b_k i) v_k = u + iv$.
Independence. If $\sum_k (a_k + b_k i) v_k = 0 + i0$, then $(*)$ and componentwise equality in $V_{\mathbf{C}} = V \times V$ give $\sum_k a_k v_k = \sum_k b_k v_k = 0$ in V; independence over $\mathbf{R}$ makes every real a_k, b_k zero.

Hence $v_1, \ldots, v_n$ is a basis of $V_{\mathbf{C}}$ over $\mathbf{C}$ by 2.26.

2.3 Exercises 2C

Problem 2C.1. Show that the subspaces of $\mathbb{R}^2$ are precisely $\{0\}$, all lines in $\mathbb{R}^2$ containing the origin, and $\mathbb{R}^2$.

Solution. A subspace of $\mathbb{R}^2$ has dimension 0, 1, or 2, and those three cases are exactly the three listed sets.
A line through the origin is precisely the span of one nonzero vector: if $L = \{w + \lambda v : \lambda \in \mathbb{R}\}$ with $v \neq 0$ and $0 \in L$, then $w = -\lambda_0 v$ for some λ_0, so

$$L = \{(\lambda - \lambda_0)v : \lambda \in \mathbb{R}\} = \operatorname{span}(v);$$

conversely $\operatorname{span}(v) = \{0 + \lambda v : \lambda \in \mathbb{R}\}$ is such a line when $v \neq 0$. Each listed set is therefore a subspace, spans being subspaces (2.6). Conversely a subspace U is finite-dimensional (2.25) with $\dim U \leq \dim \mathbb{R}^2 = 2$ by 2.37 and 2.36, and:
(i) $\dim U = 0$: the empty list is a basis, so $U = \{0\}$ by 2.4.
(ii) $\dim U = 1$: a basis of U is a single $v \neq 0$ (a list containing 0 is dependent), so $U = \operatorname{span}(v)$ is a line through the origin.
(iii) $\dim U = 2 = \dim \mathbb{R}^2$: then $U = \mathbb{R}^2$ by 2.39.

Problem 2C.2. Show that the subspaces of $\mathbb{R}^3$ are precisely $\{0\}$, all lines in $\mathbb{R}^3$ containing the origin, all planes in $\mathbb{R}^3$ containing the origin, and $\mathbb{R}^3$.

Solution. A subspace of $\mathbb{R}^3$ has dimension 0, 1, 2, or 3, and those four cases are exactly the four listed sets.
Lines through the origin are the 1*-dimensional subspaces.* If $L = \{w + \lambda v : \lambda \in \mathbb{R}\}$ with $v \neq 0$ and $0 \in L$, then $w = -\lambda_0 v$ for some λ_0, so $L = \operatorname{span}(v)$, of dimension 1 since v alone is linearly independent. Conversely a 1-dimensional subspace is $\operatorname{span}(v) = \{0 + \lambda v : \lambda \in \mathbb{R}\}$ for a basis vector $v \neq 0$.
Planes through the origin are the 2*-dimensional subspaces.* If $P = \{w + \lambda_1 v_1 + \lambda_2 v_2 : \lambda_i \in \mathbb{R}\}$ with v_1, v_2 linearly independent and $0 \in P$, then $w = -(\mu_1 v_1 + \mu_2 v_2)$ for some μ_1, μ_2, so

$$\begin{aligned} P &= \{(\lambda_1 - \mu_1)v_1 + (\lambda_2 - \mu_2)v_2 : \lambda_i \in \mathbb{R}\} \\ &= \operatorname{span}(v_1, v_2), \end{aligned}$$

a subspace (2.6) with basis v_1, v_2, so $\dim P = 2$. Conversely a 2-dimensional subspace is the span of its (linearly independent) basis v_1, v_2, hence such a plane.
Each listed set is thus a subspace; conversely a subspace U of $\mathbb{R}^3$ is finite-dimensional (2.25) with $\dim U \leq \dim \mathbb{R}^3 = 3$ by 2.37 and 2.36, and:
(i) $\dim U = 0$: the empty list is a basis, so $U = \{0\}$.
(ii) $\dim U = 1$: U is a line through the origin.
(iii) $\dim U = 2$: U is a plane through the origin.
(iv) $\dim U = 3 = \dim \mathbb{R}^3$: then $U = \mathbb{R}^3$ by 2.39.

Problem 2C.3. (a) Let $U = \{p \in \mathcal{P}_4(\mathbb{F}) : p(6) = 0\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbb{F})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbb{F})$ such that $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

Solution. (a) A basis of U is $(x-6),\ (x-6)^2,\ (x-6)^3,\ (x-6)^4$. Writing $\mathcal{P}_4 = \mathcal{P}_4(\mathbb{F})$, of dimension 5 by 2.36, the shifted powers

$$1,\ (x-6),\ (x-6)^2,\ (x-6)^3,\ (x-6)^4$$

are linearly independent (in a vanishing combination the coefficient of x^4 gives $c_4 = 0$, then x^3 gives $c_3 = 0$, and so downward), so having length $5 = \dim \mathcal{P}_4$ they are a basis of $\mathcal{P}_4$ by 2.38. Expanding p in them,

$$p = c_0 + c_1(x-6) + c_2(x-6)^2 + c_3(x-6)^3 + c_4(x-6)^4,$$

gives $p(6) = c_0$, so $p \in U$ if and only if $c_0 = 0$. The four listed polynomials thus span U and are independent as a sublist; $\dim U = 4$.
(b) Adjoin the constant polynomial 1, recovering the basis of $\mathcal{P}_4$ above.
(c) Take $W = \operatorname{span}(1)$. For $p \in \mathcal{P}_4$, let c be the constant polynomial $p(6)$; then $p = (p-c)+c$ with $p - c \in U$ and $c \in W$, giving $\mathcal{P}_4 = U + W$. If $q \in U \cap W$, then q is a constant with $q(6) = 0$, so $q = 0$; by 1.46 the sum is direct, and $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

Problem 2C.4. (a) Let $U = \{p \in \mathcal{P}_4(\mathbb{R}) : p''(6) = 0\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbb{R})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbb{R})$ such that $\mathcal{P}_4(\mathbb{R}) = U \oplus W$.

Solution. (a) A basis of U is $1,\ (x-6),\ (x-6)^3,\ (x-6)^4$. Writing $\mathcal{P}_4 = \mathcal{P}_4(\mathbb{R})$, of dimension 5 by 2.36, the shifted powers

$$1,\ (x-6),\ (x-6)^2,\ (x-6)^3,\ (x-6)^4$$

are a basis of $\mathcal{P}_4$ by 2C.3 and 2.38. Expanding $p = \sum_k c_k(x-6)^k$ and differentiating twice,

$$p'' = 2c_2 + 6c_3(x-6) + 12c_4(x-6)^2,$$

so $p''(6) = 2c_2$ and $p \in U$ if and only if $c_2 = 0$. The four listed polynomials thus span U and are independent as a sublist; $\dim U = 4$.
(b) Adjoin $(x-6)^2$, recovering the basis of $\mathcal{P}_4$ above.
(c) Take $W = \operatorname{span}\big((x-6)^2\big)$. For $p \in \mathcal{P}_4$ set $\lambda = \frac{1}{2}p''(6)$; since $\big(\lambda(x-6)^2\big)'' = 2\lambda = p''(6)$,

$$p = \big(p - \lambda(x-6)^2\big) + \lambda(x-6)^2$$

has first summand in U and second in W, so $\mathcal{P}_4 = U + W$. If $q \in U \cap W$, then $q = c(x-6)^2$ with $q''(6) = 2c = 0$, so $q = 0$; by 1.46 the sum is direct, and $\mathcal{P}_4(\mathbb{R}) = U \oplus W$.

Problem 2C.5. (a) Let $U = \{p \in \mathcal{P}_4(\mathbb{F}) : p(2) = p(5)\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbb{F})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbb{F})$ such that $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

Solution. (a) A basis of U is $1,\ (x-2)(x-5),\ x(x-2)(x-5),\ x^2(x-2)(x-5)$. Writing $\mathcal{P}_4 = \mathcal{P}_4(\mathbb{F})$, of dimension 5 by 2.36, the list

$$1, \quad x-2, \quad (x-2)(x-5), \quad x(x-2)(x-5), \quad x^2(x-2)(x-5)$$

is monic of degrees $0, 1, 2, 3, 4$, hence linearly independent (in a vanishing combination the coefficient of x^4 gives $c_4 = 0$, then x^3 gives $c_3 = 0$, and so downward), so having length $5 = \dim \mathcal{P}_4$ it is a basis of $\mathcal{P}_4$ by 2.38. Expand $p \in \mathcal{P}_4$ in it with coefficients $c_0, \dots, c_4$; every term after the second vanishes at 2 and at 5, so

$$p(2) = c_0, \qquad p(5) = c_0 + 3c_1,$$

so $p \in U$ if and only if $3c_1 = 0$, i.e. (as 3 is invertible in $\mathbb{F}$) if and only if $c_1 = 0$. Hence U is the span of the four listed polynomials, each of which does lie in U, and they are linearly independent as a sublist; $\dim U = 4$.
(b) Adjoin $x - 2$, recovering the basis of $\mathcal{P}_4$ above.
(c) Take $W = \operatorname{span}(x-2)$. Given $p \in \mathcal{P}_4$, set $\lambda = \frac{1}{3}\big(p(5) - p(2)\big)$ and $q = p - \lambda(x-2)$; then $q(2) = p(2)$ and

$$q(5) = p(5) - 3\lambda = p(2) = q(2),$$

so $p = q + \lambda(x-2)$ lies in $U + W$. If $r \in U \cap W$, then $r = c(x-2)$ and $r(2) = r(5)$ gives $0 = 3c$, so $r = 0$; by 1.46 the sum is direct: $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

Problem 2C.6. (a) Let $U = \{p \in \mathcal{P}_4(\mathbb{F}) : p(2) = p(5) = p(6)\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbb{F})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbb{F})$ such that $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

Solution. (a) A basis of U is $1,\ (x-2)(x-5)(x-6),\ x(x-2)(x-5)(x-6)$. Writing $\mathcal{P}_4 = \mathcal{P}_4(\mathbb{F})$, of dimension 5 by 2.36, the list

$$1, \quad x-2, \quad (x-2)(x-5), \quad (x-2)(x-5)(x-6), \quad x(x-2)(x-5)(x-6)$$

is monic of degrees $0, 1, 2, 3, 4$, hence linearly independent (the coefficient of x^4 gives $c_4 = 0$, then x^3 gives $c_3 = 0$, and so downward), so having length $5 = \dim \mathcal{P}_4$ it is a basis of $\mathcal{P}_4$ by 2.38. Expand $p \in \mathcal{P}_4$ in it with coefficients $c_0, \ldots, c_4$; the last two terms vanish at each of $2, 5, 6$, so

$$p(2) = c_0, \qquad p(5) = c_0 + 3c_1, \qquad p(6) = c_0 + 4c_1 + 4c_2,$$

and $p \in U$ if and only if $3c_1 = 0$ and $4c_1 + 4c_2 = 0$, i.e. (as 3 and 4 are invertible in $\mathbb{F}$) if and only if $c_1 = c_2 = 0$. Hence U is the span of the three listed polynomials, all of which lie in U (the constant takes one value everywhere, the other two vanish at $2, 5, 6$), and they are independent as a sublist; $\dim U = 3$.
(b) Adjoin $x - 2$ and $(x-2)(x-5)$, recovering after reordering the basis of $\mathcal{P}_4$ above.
(c) Take $W = \operatorname{span}\big(x-2,\ (x-2)(x-5)\big)$. The basis in (b) spans $\mathcal{P}_4$ and each of its vectors lies in U or in W, so $\mathcal{P}_4 = U + W$. If $q \in U \cap W$, then by (a) and the definition of W,

$$a_0 + a_3(x-2)(x-5)(x-6) + a_4\, x(x-2)(x-5)(x-6) = b_1(x-2) + b_2(x-2)(x-5)$$

for some scalars; subtracting writes 0 in the basis from (b), so all coefficients vanish by 2.28 and $q = 0$. By 1.46 the sum is direct: $\mathcal{P}_4(\mathbb{F}) = U \oplus W$.

Problem 2C.7. (a) Let $U = \left\{p \in \mathcal{P}_4(\mathbb{R}) : \int_{-1}^{1} p = 0\right\}$. Find a basis of U.
(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbb{R})$.
(c) Find a subspace W of $\mathcal{P}_4(\mathbb{R})$ such that $\mathcal{P}_4(\mathbb{R}) = U \oplus W$.

Solution. (a) A basis of U is $x,\ x^2 - \frac{1}{3},\ x^3,\ x^4 - \frac{1}{5}$. Writing $\mathcal{P}_4 = \mathcal{P}_4(\mathbb{R})$, of dimension 5 by 2.36, the list

$$1, \quad x, \quad x^2 - \tfrac{1}{3}, \quad x^3, \quad x^4 - \tfrac{1}{5}$$

is monic of degrees $0, 1, 2, 3, 4$, hence linearly independent (the coefficient of x^4 forces $c_4 = 0$, then x^3 forces $c_3 = 0$, and so downward), so having length $5 = \dim \mathcal{P}_4$ it is a basis of $\mathcal{P}_4$ by 2.38. Since $\int_{-1}^{1} x^k\, dx$ equals $2/(k+1)$ for k even and 0 for k odd,

$$\int_{-1}^{1} 1 = 2, \qquad \int_{-1}^{1} x = \int_{-1}^{1} x^3 = 0,$$

while $\int_{-1}^{1}(x^2 - \frac{1}{3}) = \frac{2}{3} - \frac{2}{3} = 0$ and $\int_{-1}^{1}(x^4 - \frac{1}{5}) = \frac{2}{5} - \frac{2}{5} = 0$. So expanding p in the basis with coefficients $c_0, \ldots, c_4$ gives $\int_{-1}^{1} p = 2c_0$, and $p \in U$ if and only if $c_0 = 0$. Hence U is the span of the four listed polynomials, each of which lies in U, and they are independent as a sublist; $\dim U = 4$.
(b) Adjoin the constant polynomial 1, recovering the basis of $\mathcal{P}_4$ above.
(c) Take $W = \operatorname{span}(1)$. Given $p \in \mathcal{P}_4$, set $c = \frac{1}{2}\int_{-1}^{1} p$; then $p = (p - c) + c$ with $c \in W$ and

$$\int_{-1}^{1} (p - c) = \int_{-1}^{1} p - 2c = 0,$$

so $p - c \in U$ and $\mathcal{P}_4 = U + W$. If $q \in U \cap W$, then q is a constant a with $0 = \int_{-1}^{1} a = 2a$, so $q = 0$; by 1.46 the sum is direct: $\mathcal{P}_4(\mathbb{R}) = U \oplus W$.

Problem 2C.8. Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Prove that

$$\dim \operatorname{span}(v_1 + w, \ldots, v_m + w) \geq m - 1.$$

Solution. The list $v_1 - v_m, \ldots, v_{m-1} - v_m$ is a linearly independent list of length $m - 1$ lying in $U = \operatorname{span}(v_1 + w, \ldots, v_m + w)$, which gives the bound. (For $m = 1$ the claim is only $\dim U \geq 0$, so assume $m \geq 2$.)
Each $v_k - v_m = (v_k + w) - (v_m + w)$ lies in U, and if $\sum_{k=1}^{m-1} a_k(v_k - v_m) = 0$, regrouping gives

$$a_1 v_1 + \cdots + a_{m-1} v_{m-1} - (a_1 + \cdots + a_{m-1}) v_m = 0,$$

so linear independence of $v_1, \ldots, v_m$ forces $a_1 = \cdots = a_{m-1} = 0$.
Since U is spanned by a finite list it is finite-dimensional, and a basis of U is a spanning list of U of length $\dim U$; by 2.22 the independent list above is no longer, i.e.

$$m - 1 \leq \dim U = \dim \operatorname{span}(v_1 + w, \ldots, v_m + w).$$

Problem 2C.9. Suppose m is a positive integer and $p_0, p_1, \ldots, p_m \in \mathcal{P}(\mathbb{F})$ are such that each p_k has degree k. Prove that $p_0, p_1, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbb{F})$.

Solution. Each p_k has degree $k \leq m$, so $p_0, \ldots, p_m$ is a list of $m + 1 = \dim \mathcal{P}_m(\mathbb{F})$ vectors in $\mathcal{P}_m(\mathbb{F})$ (2.36); by 2.38 only linear independence needs checking.
Suppose $a_0 p_0 + \cdots + a_m p_m = 0$ with not every a_k zero, and let k be the largest index with $a_k \neq 0$, so that

$$a_0 p_0 + a_1 p_1 + \cdots + a_k p_k = 0.$$

Writing $p_j(z) = c_{j,0} + \cdots + c_{j,j} z^j$ with $c_{j,j} \neq 0$, the terms with $j < k$ have degree $j < k$ and so contribute nothing to the coefficient of z^k on the left, which is therefore $a_k c_{k,k} \neq 0$. But the coefficients of a polynomial are uniquely determined (as noted after 4.8: a polynomial with a nonzero coefficient has only finitely many zeros, while $\mathbb{F}$ is infinite), so the left side being the zero function forces that coefficient to be 0, a contradiction.
Hence every $a_k = 0$, and $p_0, p_1, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbb{F})$.

Problem 2C.10. Suppose m is a positive integer. For $0 \leq k \leq m$, let

$$p_k(x) = x^k (1 - x)^{m-k}.$$

Show that $p_0, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbb{F})$.
[The basis in this exercise leads to what are called Bernstein polynomials. You can do a web search to learn how Bernstein polynomials are used to approximate continuous functions on $[0, 1]$.]

Solution. Each p_k has degree $k + (m - k) = m$, so $p_0, \ldots, p_m$ is a list of $m + 1 = \dim \mathcal{P}_m(\mathbb{F})$ vectors in $\mathcal{P}_m(\mathbb{F})$ (2.36); by 2.38 only linear independence needs checking.
Suppose $a_0 p_0 + \cdots + a_m p_m = 0$ with not every a_k zero, and let j be the smallest index with $a_j \neq 0$, so that $\sum_{k=j}^{m} a_k x^k (1 - x)^{m-k} = 0$. By the binomial theorem,

$$x^k (1 - x)^{m-k} = \sum_{i=0}^{m-k} \binom{m-k}{i} (-1)^i x^{k+i},$$

so every monomial of p_k has degree at least k; hence the terms with $k > j$ contribute nothing to the coefficient of x^j on the left, which is therefore $a_j \neq 0$. Since the coefficients of a polynomial are

uniquely determined (as noted after 4.8, and used in 2C.9), the left side being the zero function forces $a_j = 0$, a contradiction.
Hence every $a_k = 0$, and $p_0, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbb{F})$.

Problem 2C.11. Suppose U and W are both four-dimensional subspaces of $\mathbb{C}^6$. Prove that there exist two vectors in $U \cap W$ such that neither of these vectors is a scalar multiple of the other.

Solution. Take the first two vectors u_1, u_2 of a basis of $U \cap W$, a list of length at least 2. Indeed $U \cap W$ and $U + W$ are subspaces of $\mathbb{C}^6$ (Exercise 10 in Section 1C, and 1.40), hence finite-dimensional by 2.25 with $\dim(U + W) \le \dim \mathbb{C}^6 = 6$ by 2.37 and 2.36, so 2.43 gives

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) \ge 4 + 4 - 6 = 2.$$

A sublist of a linearly independent list is linearly independent, so u_1, u_2 is linearly independent. Were $u_2 = \lambda u_1$, then $\lambda u_1 + (-1)u_2 = 0$ would be a vanishing combination with a nonzero coefficient, and symmetrically for $u_1 = \lambda u_2$. Hence neither of the vectors $u_1, u_2 \in U \cap W$ is a scalar multiple of the other.

Problem 2C.12. Suppose that U and W are subspaces of $\mathbb{R}^8$ such that $\dim U = 3$, $\dim W = 5$, and $U + W = \mathbb{R}^8$. Prove that $\mathbb{R}^8 = U \oplus W$.

Solution. The intersection $U \cap W$ is trivial, so 1.46 makes the sum direct. Indeed $U \cap W$ is a subspace of $\mathbb{R}^8$ (Exercise 10 in Section 1C), hence finite-dimensional by 2.25, and since $\dim(U + W) = \dim \mathbb{R}^8 = 8$ by hypothesis and 2.36, applying 2.43 to U and W gives

$$8 = 3 + 5 - \dim(U \cap W), \quad \text{so} \quad \dim(U \cap W) = 0.$$

A space of dimension 0 has the empty list as a basis, whose span is $\{0\}$; thus $U \cap W = \{0\}$ and $\mathbb{R}^8 = U + W = U \oplus W$.

Problem 2C.13. Suppose U and W are both five-dimensional subspaces of $\mathbb{R}^9$. Prove that $U \cap W \neq \{0\}$.

Solution. The intersection has dimension at least 1. Both $U + W$ and $U \cap W$ are subspaces of $\mathbb{R}^9$ (1.40 and Exercise 10 in Section 1C), hence finite-dimensional by 2.25 with $\dim(U + W) \le \dim \mathbb{R}^9 = 9$ by 2.37 and 2.36, so 2.43 gives

$$\dim(U \cap W) = \dim U + \dim W - \dim(U + W) \ge 5 + 5 - 9 = 1.$$

Since $\dim\{0\} = 0$ (the empty list being a basis), $U \cap W \neq \{0\}$.

Problem 2C.14. Suppose V is a ten-dimensional vector space and V_1, V_2, V_3 are subspaces of V with $\dim V_1 = \dim V_2 = \dim V_3 = 7$. Prove that $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Solution. Two applications of 2.43 force $\dim(V_1 \cap V_2 \cap V_3) \ge 1$. All the subspaces below are subspaces of V (Exercise 10 in Section 1C for intersections, 1.40 for sums), hence finite-dimensional by 2.25 with dimension at most $\dim V = 10$ by 2.37. So

$$\begin{aligned}\dim(V_1 \cap V_2) &= \dim V_1 + \dim V_2 - \dim(V_1 + V_2)\\ &\ge 7 + 7 - 10 = 4,\end{aligned}$$

and, since $(V_1 \cap V_2) \cap V_3 = V_1 \cap V_2 \cap V_3$,

$$\begin{aligned}\dim(V_1 \cap V_2 \cap V_3) &= \dim(V_1 \cap V_2) + \dim V_3 - \dim\big((V_1 \cap V_2) + V_3\big) \\ &\geq 4 + 7 - 10 = 1.\end{aligned}$$

Since $\dim\{0\} = 0$ (the empty list is a basis of $\{0\}$), $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Problem 2C.15. Suppose V is finite-dimensional and V_1, V_2, V_3 are subspaces of V with $\dim V_1 + \dim V_2 + \dim V_3 > 2 \dim V$. Prove that $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Solution. For any subspaces A, B of V (all finite-dimensional by 2.25),

$$\dim(A \cap B) = \dim A + \dim B - \dim(A + B) \geq \dim A + \dim B - \dim V,$$

by 2.43 together with $\dim(A + B) \leq \dim V$, which holds by 2.37 since $A + B$ is a subspace of V (1.40). Applying this with $A = V_1, B = V_2$, then with $A = V_1 \cap V_2, B = V_3$,

$$\begin{aligned}\dim(V_1 \cap V_2 \cap V_3) &\geq \dim(V_1 \cap V_2) + \dim V_3 - \dim V \\ &\geq \dim V_1 + \dim V_2 + \dim V_3 - 2 \dim V,\end{aligned}$$

which is positive by hypothesis. Since $\dim\{0\} = 0$ (the empty list is its basis), $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Problem 2C.16. Suppose V is finite-dimensional and U is a subspace of V with $U \neq V$. Let $n = \dim V$ and $m = \dim U$. Prove that there exist $n - m$ subspaces of V, each of dimension $n - 1$, whose intersection equals U.

Solution. Extend a basis $u_1, \ldots, u_m$ of U to a basis $u_1, \ldots, u_m, v_1, \ldots, v_{n-m}$ of V (2.32; the extension has length $\dim V = n$ by 2.34) and take, for $k \in \{1, \ldots, n - m\}$,

$$\begin{aligned}W_k = \operatorname{span}(&u_1, \ldots, u_m,\ v_1, \ldots, v_{k-1}, \\ &v_{k+1}, \ldots, v_{n-m}),\end{aligned}$$

the span of that basis with v_k deleted. There are indeed $n - m \geq 1$ of these, since $m \leq n$ by 2.37 while $m = n$ would give $U = V$ by 2.39.
Each such spanning list has length $n - 1$ and is linearly independent, being a sublist of a basis; so it is a basis of W_k and $\dim W_k = n - 1$.
Since $u_1, \ldots, u_m$ occurs in every spanning list, $U \subseteq W_1 \cap \cdots \cap W_{n-m}$. Conversely, let v lie in that intersection and write $v = \sum_j a_j u_j + \sum_k b_k v_k$ in the basis of V. Fixing k, membership in W_k also expresses v in that same basis with coefficient 0 on v_k, so $b_k = 0$ by the uniqueness in 2.28; as k was arbitrary, $v = \sum_j a_j u_j \in U$. Hence

$$W_1 \cap \cdots \cap W_{n-m} = U.$$

Problem 2C.17. Suppose that $V_1, \ldots, V_m$ are finite-dimensional subspaces of V. Prove that $V_1 + \cdots + V_m$ is finite-dimensional and

$$\dim(V_1 + \cdots + V_m) \leq \dim V_1 + \cdots + \dim V_m.$$

[The inequality above is an equality if and only if $V_1 + \cdots + V_m$ is a direct sum, as will be shown in 3.94.]

Solution. Concatenating bases does it. Put $n_k = \dim V_k$, choose a basis $v_{k,1}, \ldots, v_{k,n_k}$ of each V_k (2.31), and let L be the concatenated list

$$v_{1,1}, \ldots, v_{1,n_1},\ v_{2,1}, \ldots, v_{2,n_2},$$
$$\ldots,\ v_{m,1}, \ldots, v_{m,n_m},$$

of length $n_1 + \cdots + n_m$.
L spans $V_1 + \cdots + V_m$: every element of the sum is $u_1 + \cdots + u_m$ with $u_k \in V_k$ (1.36), and expanding each u_k in its basis exhibits it as a linear combination of vectors of L; conversely each vector of L lies in some V_k, hence in the subspace $V_1 + \cdots + V_m$ (1.40).
So $V_1 + \cdots + V_m$ has a finite spanning list and is therefore finite-dimensional; any basis of it is a linearly independent list in it, so by 2.22 that basis is no longer than L:

$$\dim(V_1 + \cdots + V_m) \leq n_1 + \cdots + n_m = \dim V_1 + \cdots + \dim V_m.$$

Problem 2C.18. Suppose V is finite-dimensional, with $\dim V = n \geq 1$. Prove that there exist one-dimensional subspaces $V_1, \ldots, V_n$ of V such that

$$V = V_1 \oplus \cdots \oplus V_n.$$

Solution. Take $V_k = \operatorname{span}(v_k)$, where $v_1, \ldots, v_n$ is a basis of V (2.31), of length $\dim V = n$.
Each V_k is one-dimensional. No v_k is 0, since a list containing 0 is dependent; and a one-vector list $v_k \neq 0$ is linearly independent, as $av_k = 0$ with $a \neq 0$ gives $v_k = a^{-1}(av_k) = 0$. So v_k is a basis of V_k.
The sum is V. Each $v \in V$ is $v = a_1 v_1 + \cdots + a_n v_n$ with $a_k v_k \in V_k$, so $V \subseteq V_1 + \cdots + V_n$, and the reverse inclusion holds since each V_k is a subspace of V.
The sum is direct. If $w_1 + \cdots + w_n = 0$ with $w_k = a_k v_k \in V_k$, then

$$a_1 v_1 + \cdots + a_n v_n = 0,$$

so every $a_k = 0$ by linear independence, whence every $w_k = 0$; apply 1.45.
Hence $V = V_1 \oplus \cdots \oplus V_n$ with each V_k one-dimensional.

Problem 2C.19. Explain why you might guess, motivated by analogy with the formula for the number of elements in the union of three finite sets, that if V_1, V_2, V_3 are subspaces of a finite-dimensional vector space, then

$$\begin{aligned}\dim(V_1 + V_2 + V_3) = {} & \dim V_1 + \dim V_2 + \dim V_3 \\ & - \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3) \\ & + \dim(V_1 \cap V_2 \cap V_3).\end{aligned}$$

Then either prove the formula above or give a counterexample.

Solution. The formula is false, the three lines

$$V_1 = \operatorname{span}\big((1,0)\big),$$
$$V_2 = \operatorname{span}\big((0,1)\big),$$
$$V_3 = \operatorname{span}\big((1,1)\big)$$

in $\mathbf{R}^2$ being a counterexample.
Why one might guess it. The table preceding these exercises matches a finite set S with a space V, $\#S$ with $\dim V$, union with sum (each the smallest object containing the two, by 1.40), and intersection with intersection. Under that dictionary two-set inclusion-exclusion, $\#(S_1 \cup S_2) = \#S_1 + \#S_2 - \#(S_1 \cap S_2)$, becomes the true theorem 2.43; translating the three-set counting formula term by term gives the displayed guess.
The guess is false. Each V_i above is the span of one nonzero vector, so $\dim V_i = 1$. Their sum contains $(1,0)$ and $(0,1)$, hence equals $\mathbf{R}^2$, and the left side is 2. Every pairwise intersection is $\{0\}$: a vector of

$V_1 \cap V_3$ is both $(a, 0)$ and (b, b), forcing $b = 0$, and comparing the vanishing coordinate likewise kills $V_1 \cap V_2$ (where $(a, 0) = (0, b)$) and $V_2 \cap V_3$ (where $(0, a) = (b, b)$). So $V_1 \cap V_2 \cap V_3 = \{0\}$ as well, all four intersection terms have dimension 0, and the right side is

$$1 + 1 + 1 - 0 - 0 - 0 + 0 = 3 \neq 2.$$

Problem 2C.20. Prove that if V_1, V_2, and V_3 are subspaces of a finite-dimensional vector space, then

$$\begin{aligned}\dim(V_1 + V_2 + V_3) = {} & \dim V_1 + \dim V_2 + \dim V_3 \\ & - \frac{\dim(V_1 \cap V_2) + \dim(V_1 \cap V_3) + \dim(V_2 \cap V_3)}{3} \\ & - \frac{\dim\big((V_1 + V_2) \cap V_3\big) + \dim\big((V_1 + V_3) \cap V_2\big) + \dim\big((V_2 + V_3) \cap V_1\big)}{3}.\end{aligned}$$

[The formula above may seem strange because the right side does not look like an integer.]

Solution. Average the three ways of grouping the sum. Every space below is a subspace of the ambient finite-dimensional space (1.40 for sums, 1.34 for intersections) and so is finite-dimensional by 2.25, letting 2.43 be applied freely. Applying it first to $V_1 + V_2$ and V_3, then to V_1 and V_2,

$$\begin{aligned}\dim(V_1 + V_2 + V_3) &= \dim(V_1 + V_2) + \dim V_3 - \dim\big((V_1 + V_2) \cap V_3\big) \\ &= \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 \cap V_2) \\ &\quad - \dim\big((V_1 + V_2) \cap V_3\big).\end{aligned}$$

Call this identity (1). Since addition of subspaces is commutative and associative, the same computation applied to the groupings $(V_1 + V_3) + V_2$ and $(V_2 + V_3) + V_1$ gives identities (2) and (3), namely (1) with the index swaps $2 \leftrightarrow 3$ and $1 \leftrightarrow 3$; their subtracted terms are $\dim(V_1 \cap V_3)$, $\dim((V_1 + V_3) \cap V_2)$ and $\dim(V_2 \cap V_3)$, $\dim((V_2 + V_3) \cap V_1)$.

All three identities have left side $\dim(V_1 + V_2 + V_3)$, so adding them and dividing by 3 leaves that side unchanged while averaging the right sides: the three pairwise-intersection terms become the first displayed fraction of the statement and the three sum-intersection terms the second. That is the claimed formula.

3 Linear Maps

Contents

3.1 Exercises 3A

Problem 3A.1. Suppose $b, c \in \mathbb{R}$. Define $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ by

$$T(x,y,z) = (2x - 4y + 3z + b,\ 6x + cxyz).$$

Show that T is linear if and only if $b = c = 0$.

Solution. If $b = c = 0$, then $T(x,y,z) = (2x - 4y + 3z,\ 6x)$ satisfies the two conditions in 3.1: for $u = (x_1, y_1, z_1)$, $v = (x_2, y_2, z_2)$ and $\lambda \in \mathbb{R}$, regrouping coordinatewise gives

$$\begin{aligned} T(u+v) &= \bigl(2(x_1+x_2) - 4(y_1+y_2) + 3(z_1+z_2),\ 6(x_1+x_2)\bigr) \\ &= (2x_1 - 4y_1 + 3z_1,\ 6x_1) + (2x_2 - 4y_2 + 3z_2,\ 6x_2) = Tu + Tv, \\ T(\lambda v) &= (2\lambda x - 4\lambda y + 3\lambda z,\ 6\lambda x) = \lambda(Tv). \end{aligned}$$

Conversely, suppose T is linear. Then $T(0,0,0) = (0,0)$ by 3.10, while the defining formula gives $T(0,0,0) = (b,0)$; comparing first coordinates, $b = 0$. For c, homogeneity with $\lambda = 2$ and $v = (1,1,1)$ compares the second coordinates of

$$T(2,2,2) = (2 + b,\ 12 + 8c), \qquad 2\,T(1,1,1) = (2 + 2b,\ 12 + 2c),$$

since $c \cdot 2 \cdot 2 \cdot 2 = 8c$; thus $12 + 8c = 12 + 2c$ and $c = 0$.

Problem 3A.2. Suppose $b, c \in \mathbb{R}$. Define $T\colon \mathcal{P}(\mathbb{R}) \to \mathbb{R}^2$ by

$$Tp = \left(3p(4) + 5p'(6) + b\,p(1)p(2),\ \int_{-1}^{2} x^3 p(x)\,dx + c\sin p(0)\right).$$

Show that T is linear if and only if $b = c = 0$.

Solution. If $b = c = 0$, then

$$Tp = \left(3p(4) + 5p'(6),\ \int_{-1}^{2} x^3 p(x)\,dx\right)$$

satisfies the two conditions in 3.1: evaluation at a point, differentiation and the integral are each additive and homogeneous, so every coordinate of $T(p+q)$ is the sum of the corresponding coordinates of Tp and Tq, and every coordinate of $T(\lambda p)$ is λ times that of Tp, addition and scalar multiplication in $\mathbb{R}^2$ being coordinatewise.

Conversely, suppose T is linear. Take p the constant polynomial 1 (so $p' = 0$); the first coordinates of Tp and of $T(2p)$ are $3 + b$ and $3 \cdot 2 + b \cdot 2 \cdot 2 = 6 + 4b$, so homogeneity with $\lambda = 2$ gives

$$6 + 4b = 2(3 + b) = 6 + 2b,$$

whence $b = 0$. Now take q the constant polynomial $\pi/2$, and note $\int_{-1}^{2} x^3\,dx = \frac{16}{4} - \frac{1}{4} = \frac{15}{4}$. Since $\sin(\pi/2) = 1$ and $\sin\pi = 0$, the second coordinates of Tq and of $T(2q)$ are $\frac{15\pi}{8} + c$ and $\frac{15\pi}{4}$, so homogeneity with $\lambda = 2$ gives

$$\frac{15\pi}{4} = 2\left(\frac{15\pi}{8} + c\right) = \frac{15\pi}{4} + 2c,$$

whence $c = 0$.

Problem 3A.3. Suppose that $T \in \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$. Show that there exist scalars $A_{j,k} \in \mathbb{F}$ for $j = 1, \ldots, m$ and $k = 1, \ldots, n$ such that

$$T(x_1, \ldots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n,\ \ldots,\ A_{m,1}x_1 + \cdots + A_{m,n}x_n)$$

for every $(x_1, \dots, x_n) \in \mathbb{F}^n$.
[This exercise shows that the linear map T has the form promised in the second to last item of Example 3.3.]

Solution. Take $A_{j,k}$ to be the j^{th} coordinate of Te_k, where $e_1, \dots, e_n$ is the standard basis of $\mathbb{F}^n$ (2.27); that is,

$$Te_k = (A_{1,k},\ A_{2,k},\ \dots,\ A_{m,k}) \qquad (k = 1, \dots, n).$$

Since $(x_1, \dots, x_n) = x_1 e_1 + \cdots + x_n e_n$, linearity of T and the coordinatewise vector space operations on $\mathbb{F}^m$ give

$$\begin{aligned} T(x_1, \dots, x_n) &= x_1(Te_1) + \cdots + x_n(Te_n) \\ &= x_1(A_{1,1}, \dots, A_{m,1}) + \cdots + x_n(A_{1,n}, \dots, A_{m,n}) \\ &= (A_{1,1}x_1 + \cdots + A_{1,n}x_n,\ \dots,\ A_{m,1}x_1 + \cdots + A_{m,n}x_n), \end{aligned}$$

the last step because the j^{th} coordinate of the middle expression is $A_{j,1}x_1 + \cdots + A_{j,n}x_n$.

Problem 3A.4. Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_m$ is a list of vectors in V such that $Tv_1, \dots, Tv_m$ is a linearly independent list in W. Prove that $v_1, \dots, v_m$ is linearly independent.

Solution. Suppose $a_1, \dots, a_m \in \mathbb{F}$ satisfy $a_1 v_1 + \cdots + a_m v_m = 0$. Applying T and using linearity, with $T0 = 0$ by 3.10,

$$a_1(Tv_1) + \cdots + a_m(Tv_m) = T(a_1 v_1 + \cdots + a_m v_m) = T0 = 0.$$

Linear independence of $Tv_1, \dots, Tv_m$ now forces $a_1 = \cdots = a_m = 0$. Hence $v_1, \dots, v_m$ is linearly independent.

Problem 3A.5. Prove that $\mathcal{L}(V, W)$ is a vector space, as was asserted in 3.6.

Solution. Every condition in 1.20 follows by evaluating both sides at an arbitrary $v \in V$ and quoting the corresponding property of W, the operations being $(S + T)(v) = Sv + Tv$ and $(\lambda T)(v) = \lambda(Tv)$ (3.5).
Closure, for $u, v \in V$ and $\mu \in \mathbb{F}$:

$$\begin{aligned} (S+T)(u+v) &= (Su+Sv) + (Tu+Tv) = (S+T)(u) + (S+T)(v), \\ (S+T)(\mu v) &= \mu(Sv) + \mu(Tv) = \mu\big((S+T)(v)\big), \\ (\lambda T)(u+v) &= \lambda(Tu) + \lambda(Tv) = (\lambda T)(u) + (\lambda T)(v), \\ (\lambda T)(\mu v) &= (\lambda\mu)(Tv) = (\mu\lambda)(Tv) = \mu((\lambda T)v), \end{aligned}$$

using additivity and homogeneity of S, T, distributivity and associativity of scalar multiplication in W, and commutativity in $\mathbb{F}$; so $S + T, \lambda T \in \mathcal{L}(V, W)$. The remaining axioms, evaluated at v:

$$\begin{aligned} (S+T)(v) &= Sv + Tv = Tv + Sv = (T+S)(v), \\ \big((R+S)+T\big)(v) &= (Rv+Sv) + Tv = Rv + (Sv+Tv) = \big(R+(S+T)\big)(v), \\ ((\lambda\mu)T)v &= \lambda\big(\mu(Tv)\big) = \big(\lambda(\mu T)\big)v, \\ (T+0)(v) &= Tv + 0 = Tv, \\ \big(T+(-1)T\big)(v) &= Tv + (-1)(Tv) = 0, \\ (1T)(v) &= 1(Tv) = Tv, \\ \big(\lambda(S+T)\big)(v) &= \lambda(Sv) + \lambda(Tv) = (\lambda S + \lambda T)(v), \\ \big((\lambda+\mu)T\big)(v) &= \lambda(Tv) + \mu(Tv) = (\lambda T + \mu T)(v), \end{aligned}$$

where the additive identity is the zero linear map of Example 3.3, the additive inverse $-T = (-1)T$ works by 1.32 applied in W, and each right-hand step is the matching axiom of W or of $\mathbb{F}$. Hence $\mathcal{L}(V, W)$ is a vector space over $\mathbb{F}$.

Problem 3A.6. Prove that multiplication of linear maps has the associative, identity, and distributive properties asserted in 3.8.

Solution. Each identity follows by evaluating both sides at a point of the common domain, using $(ST)(u) = S(Tu)$ from 3.7.
Associativity: with $T_3 \in \mathcal{L}(V_1, V_2)$, $T_2 \in \mathcal{L}(V_2, V_3)$, $T_1 \in \mathcal{L}(V_3, V_4)$, both sides are maps $V_1 \to V_4$, and for $u \in V_1$,

$$\big((T_1T_2)T_3\big)(u) = T_1\big(T_2(T_3u)\big) = \big(T_1(T_2T_3)\big)(u).$$

Identity: for $T \in \mathcal{L}(V, W)$, with I the identity operator on V in TI and on W in IT,

$$(TI)(v) = T(Iv) = Tv = I(Tv) = (IT)(v) \qquad (v \in V).$$

Distributivity: for $T, T_1, T_2 \in \mathcal{L}(U, V)$, $S, S_1, S_2 \in \mathcal{L}(V, W)$, and $u \in U$,

$$\begin{aligned}\big((S_1 + S_2)T\big)(u) &= S_1(Tu) + S_2(Tu) = (S_1T + S_2T)(u),\\ \big(S(T_1 + T_2)\big)(u) &= S(T_1u + T_2u) = S(T_1u) + S(T_2u) = (ST_1 + ST_2)(u),\end{aligned}$$

the second line using additivity of S.

Problem 3A.7. Show that every linear map from a one-dimensional vector space to itself is multiplication by some scalar. More precisely, prove that if $\dim V = 1$ and $T \in \mathcal{L}(V)$, then there exists $\lambda \in \mathbb{F}$ such that $Tv = \lambda v$ for all $v \in V$.

Solution. Take $\lambda \in \mathbb{F}$ with $Tu = \lambda u$, where u is a basis of the one-dimensional space V; such a λ exists because $Tu \in V = \operatorname{span}(u)$. Every $v \in V$ has the form $v = au$ with $a \in \mathbb{F}$, so homogeneity of T gives

$$Tv = T(au) = a(Tu) = a(\lambda u) = \lambda(au) = \lambda v.$$

Problem 3A.8. Give an example of a function $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ such that

$$\varphi(av) = a\varphi(v)$$

for all $a \in \mathbb{R}$ and all $v \in \mathbb{R}^2$ but φ is not linear.
[This exercise and the next exercise show that neither homogeneity nor additivity alone is enough to imply that a function is a linear map.]

Solution. Define $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ by

$$\varphi(x, y) = \begin{cases} \dfrac{x^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0),\\ 0 & \text{if } (x, y) = (0, 0).\end{cases}$$

Homogeneity: if $a = 0$ or $v = (0, 0)$ then $av = (0, 0)$, so both $\varphi(av)$ and $a\varphi(v)$ equal 0; otherwise $av = (ax, ay) \neq (0, 0)$ and

$$\varphi(av) = \frac{a^3x^3}{a^2(x^2 + y^2)} = a \cdot \frac{x^3}{x^2 + y^2} = a\varphi(v).$$

Additivity fails at $u = (1, 0)$, $v = (0, 1)$:

$$\varphi(u + v) = \varphi(1, 1) = \tfrac{1}{2} \neq 1 + 0 = \varphi(u) + \varphi(v),$$

so φ is not linear by 3.1.

Problem 3A.9. Give an example of a function $\varphi\colon \mathbb{C} \to \mathbb{C}$ such that

$$\varphi(w + z) = \varphi(w) + \varphi(z)$$

for all $w, z \in \mathbb{C}$ but φ is not linear. (Here $\mathbb{C}$ is thought of as a complex vector space.)
[There also exists a function $\varphi\colon \mathbb{R} \to \mathbb{R}$ such that φ satisfies the additivity condition above but φ is not linear. However, showing the existence of such a function involves considerably more advanced tools.]

Solution. Define $\varphi\colon \mathbb{C} \to \mathbb{C}$ to be complex conjugation:

$$\varphi(z) = \overline{z},$$

so that $\varphi(a + bi) = a - bi$ for all $a, b \in \mathbb{R}$.
Additivity: for $w = a + bi$ and $z = c + di$ with $a, b, c, d \in \mathbb{R}$,

$$\varphi(w + z) = (a + c) - (b + d)i = (a - bi) + (c - di) = \varphi(w) + \varphi(z).$$

Homogeneity fails because the scalars here include i: taking $\lambda = i$ and $z = 1$,

$$\varphi(\lambda z) = \overline{i} = -i \neq i = i\,\overline{1} = \lambda\varphi(z),$$

so φ is not linear by 3.1.

Problem 3A.10. Prove or give a counterexample: If $q \in \mathcal{P}(\mathbb{R})$ and $T\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ is defined by $Tp = q \circ p$, then T is a linear map.
[The function T defined here differs from the function T defined in the last bullet point of 3.3 by the order of the functions in the compositions.]

Solution. False: take $q(x) = x^2$, so that $(Tp)(x) = (p(x))^2$, and let p be the constant polynomial 1. Then $T(2p)$ is the constant polynomial 4 while $2(Tp)$ is the constant polynomial 2:

$$\begin{aligned}\big(T(2p)\big)(x) &= \big((2p)(x)\big)^2 = 4,\\ \big(2(Tp)\big)(x) &= 2\big(p(x)\big)^2 = 2.\end{aligned}$$

So homogeneity fails, and T is not linear by 3.1.

Problem 3A.11. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST = TS$ for every $S \in \mathcal{L}(V)$.

Solution. ($\Rightarrow$) If $T = \lambda I$, then for all $S \in \mathcal{L}(V)$ and $v \in V$, homogeneity of S gives

$$(ST)v = S(\lambda v) = \lambda(Sv) = (TS)v.$$

($\Leftarrow$) Suppose $ST = TS$ for every $S \in \mathcal{L}(V)$; we may assume $V \neq \{0\}$, since otherwise $T = 0 = 0 \cdot I$. Fix $v \neq 0$. If $Tv \notin \operatorname{span}(v)$, then v, Tv is linearly independent (in a relation $av + b(Tv) = 0$, $b \neq 0$ would put Tv in $\operatorname{span}(v)$, and then $av = 0$ with $v \neq 0$ forces $a = 0$), so by 2.32 it extends to a basis $v, Tv, u_3, \dots, u_n$ of the finite-dimensional space V, and 3.4 supplies $S \in \mathcal{L}(V)$ with $Sv = v$ and $S(Tv) = Su_3 = \cdots = Su_n = 0$, whence

$$(ST)v = S(Tv) = 0 \neq Tv = T(Sv) = (TS)v,$$

since $Tv \neq 0$ as a member of a linearly independent list. This contradiction gives $Tv = a_v v$ for a scalar $a_v \in \mathbf{F}$, unique because $v \neq 0$. For nonzero v, w the scalar a_v agrees with a_w:

(i) v, w linearly dependent, say $w = cv$ with $c \neq 0$:

$$a_w w = Tw = cTv = c\,a_v v = a_v w, \qquad \text{so } a_w = a_v.$$

(ii) v, w linearly independent, so $v + w \neq 0$:

$$a_{v+w}v + a_{v+w}w = T(v + w) = Tv + Tw = a_v v + a_w w,$$

and independence of v, w forces $a_v = a_{v+w} = a_w$.
Writing λ for this common value, $Tv = \lambda v$ for every nonzero $v \in V$, while $T0 = 0 = \lambda \cdot 0$ by 3.10. Hence $T = \lambda I$.

Problem 3A.12. Suppose U is a subspace of V with $U \neq V$. Suppose $S \in \mathcal{L}(U, W)$ and $S \neq 0$ (which means that $Su \neq 0$ for some $u \in U$). Define $T: V \to W$ by

$$Tv = \begin{cases} Sv & \text{if } v \in U, \\ 0 & \text{if } v \in V \text{ and } v \notin U. \end{cases}$$

Prove that T is not a linear map on V.

Solution. Additivity fails at the pair u, w, where $u \in U$ satisfies $Su \neq 0$ (available since $S \neq 0$) and $w \in V$ satisfies $w \notin U$ (available since $U \neq V$). Indeed $u + w \notin U$, since $u + w \in U$ would give $w = (u + w) - u \in U$ by closure of the subspace U under addition and additive inverses. Hence $T(u + w) = 0$, while $Tu = Su$ and $Tw = 0$, so

$$T(u + w) = 0 \neq Su = Tu + Tw.$$

Therefore T is not a linear map by 3.1.

Problem 3A.13. Suppose V is finite-dimensional. Prove that every linear map on a subspace of V can be extended to a linear map on V. In other words, show that if U is a subspace of V and $S \in \mathcal{L}(U, W)$, then there exists $T \in \mathcal{L}(V, W)$ such that $Tu = Su$ for all $u \in U$.
[The result in this exercise is used in the proof of 3.125.]

Solution. Take $T \in \mathcal{L}(V, W)$ to be the map furnished by the linear map lemma 3.4 on an extended basis: U is finite-dimensional by 2.25 and so has a basis $u_1, \dots, u_m$ (2.31, the empty list if $U = \{0\}$), which is linearly independent in V and hence extends by 2.32 to a basis $u_1, \dots, u_m, v_1, \dots, v_n$ of V; define

$$Tu_k = Su_k \quad (k = 1, \dots, m), \qquad Tv_j = 0 \quad (j = 1, \dots, n).$$

Every $u \in U$ is $u = c_1u_1 + \cdots + c_mu_m$ with $c_k \in \mathbf{F}$, so linearity of T and then of S give

$$Tu = c_1(Tu_1) + \cdots + c_m(Tu_m) = c_1(Su_1) + \cdots + c_m(Su_m) = Su,$$

which for $m = 0$ reads $T0 = 0 = S0$ (3.10). Hence T extends S.

Problem 3A.14. Suppose V is finite-dimensional with $\dim V > 0$, and suppose W is infinite-dimensional. Prove that $\mathcal{L}(V, W)$ is infinite-dimensional.

Solution. We exhibit a linearly independent list of arbitrary length m in the vector space $\mathcal{L}(V, W)$ (3.6); by 2.22 no finite list can then span it, so it is infinite-dimensional (2.13).
First, W contains a linearly independent list $w_1, \dots, w_m$: starting from the empty list and given a linearly independent $w_1, \dots, w_k$, we have $\operatorname{span}(w_1, \dots, w_k) \neq W$ (else this finite list would span W, contradicting infinite-dimensionality), and any w_{k+1} chosen outside that span keeps the list independent, since

in a relation $a_1w_1 + \cdots + a_{k+1}w_{k+1} = 0$ we must have $a_{k+1} = 0$ (otherwise $w_{k+1} \in \operatorname{span}(w_1, \ldots, w_k)$) and then $a_1 = \cdots = a_k = 0$.
Next, let $v_1, \ldots, v_p$ be a basis of V (2.31), with $p \geq 1$ since $\dim V > 0$, and use the linear map lemma 3.4 to define $T_1, \ldots, T_m \in \mathcal{L}(V, W)$ by

$$T_jv_1 = w_j \qquad \text{and} \qquad T_jv_k = 0 \ \text{ for } k = 2, \ldots, p.$$

If $a_1T_1 + \cdots + a_mT_m = 0$, then evaluating at v_1 using the operations 3.5 gives

$$a_1w_1 + \cdots + a_mw_m = 0,$$

so $a_1 = \cdots = a_m = 0$ and $T_1, \ldots, T_m$ is linearly independent, as required.

Problem 3A.15. Suppose $v_1, \ldots, v_m$ is a linearly dependent list of vectors in V. Suppose also that $W \neq \{0\}$. Prove that there exist $w_1, \ldots, w_m \in W$ such that no $T \in \mathcal{L}(V, W)$ satisfies $Tv_k = w_k$ for each $k = 1, \ldots, m$.

Solution. Take $w_j = w$ and $w_k = 0$ for $k \neq j$, where $w \in W$ is nonzero (available since $W \neq \{0\}$) and j is an index with $a_j \neq 0$ in a dependence relation with scalars $a_1, \ldots, a_m \in \mathbf{F}$ not all zero,

$$a_1v_1 + \cdots + a_mv_m = 0.$$

If some $T \in \mathcal{L}(V, W)$ satisfied $Tv_k = w_k$ for every k, then applying T to this relation and using $T0 = 0$ (3.10) together with $w_k = 0$ for $k \neq j$ would give

$$0 = T(a_1v_1 + \cdots + a_mv_m) = a_1w_1 + \cdots + a_mw_m = a_jw,$$

forcing $w = 0$ since $a_j \neq 0$, a contradiction. So no such T exists.

Problem 3A.16. Suppose V is finite-dimensional with $\dim V > 1$. Prove that there exist $S, T \in \mathcal{L}(V)$ such that $ST \neq TS$.

Solution. Fix a basis $v_1, \ldots, v_n$ of V, where $n = \dim V \geq 2$ (so that v_2 is available), and let the linear map lemma 3.4 supply $S, T \in \mathcal{L}(V)$ with

$$\begin{aligned} Sv_1 &= v_1, \quad Sv_k = 0 \quad \text{for } k = 2, \ldots, n, \\ Tv_1 &= v_2, \quad Tv_k = 0 \quad \text{for } k = 2, \ldots, n. \end{aligned}$$

Evaluating both compositions at v_1,

$$(ST)v_1 = Sv_2 = 0 \neq v_2 = Tv_1 = (TS)v_1,$$

where $v_2 \neq 0$ as a member of a linearly independent list. Hence $ST \neq TS$.

Problem 3A.17. Suppose V is finite-dimensional. Show that the only two-sided ideals of $\mathcal{L}(V)$ are $\{0\}$ and $\mathcal{L}(V)$.
[A subspace $\mathcal{E}$ of $\mathcal{L}(V)$ is called a **two-sided ideal** of $\mathcal{L}(V)$ if $TE \in \mathcal{E}$ and $ET \in \mathcal{E}$ for all $E \in \mathcal{E}$ and all $T \in \mathcal{L}(V)$.]

Solution. It suffices to put I inside any two-sided ideal $\mathcal{E} \neq \{0\}$: then every $T \in \mathcal{L}(V)$ equals TI by the identity property in 3.8 and lies in $\mathcal{E}$ by the ideal property, so $\mathcal{E} = \mathcal{L}(V)$. (That $\{0\}$ and $\mathcal{L}(V)$ are two-sided ideals is immediate from $T0 = 0T = 0$.)
Fix $E \in \mathcal{E}$ with $E \neq 0$ and $u \in V$ with $Eu \neq 0$. Then $V \neq \{0\}$, so V has a basis $v_1, \ldots, v_n$ with $n = \dim V \geq 1$ finite (2.31), and the linearly independent one-term list Eu extends by 2.32 to a basis

$u_1, \ldots, u_n$ of V with $u_1 = Eu$ (of length n, every basis of V having length $\dim V$). For each j, the linear map lemma 3.4 defines $A_j, B_j \in \mathcal{L}(V)$ by

$$\begin{aligned} A_j v_j &= u, \quad & A_j v_i &= 0 \ \text{ for } i \neq j, \\ B_j u_1 &= v_j, \quad & B_j u_i &= 0 \ \text{ for } i = 2, \ldots, n. \end{aligned}$$

Two applications of the ideal property give $B_j E A_j \in \mathcal{E}$ (unambiguous by associativity, 3.8), and on the basis, using $E0 = 0$ (3.10),

$$\begin{aligned} (B_j E A_j) v_i &= 0 \qquad (i \neq j), \\ (B_j E A_j) v_j &= B_j(Eu) = B_j u_1 = v_j. \end{aligned}$$

Since $\mathcal{E}$ is a subspace and V is finite-dimensional, the finite sum $R := B_1 E A_1 + \cdots + B_n E A_n$ lies in $\mathcal{E}$, and the displayed values give $R v_i = v_i$ for each i. So R and I agree on a basis, whence $R = I$ by the uniqueness half of 3.4, and $I \in \mathcal{E}$.

3.2 Exercises 3B

Problem 3B.1. Give an example of a linear map T with $\dim \operatorname{null} T = 3$ and $\dim \operatorname{range} T = 2$.

Solution. Take the linear map (Check!) $T \in \mathcal{L}(\mathbf{F}^5, \mathbf{F}^2)$ given by

$$T(z_1, z_2, z_3, z_4, z_5) = (z_1, z_2).$$

Then $Tz = 0$ exactly when $z_1 = z_2 = 0$, so $\operatorname{null} T = \{(0, 0, z_3, z_4, z_5)\}$ has basis $(0, 0, 1, 0, 0), (0, 0, 0, 1, 0), (0, 0, 0, 0, 1)$ and $\dim \operatorname{null} T = 3$; and $T(w_1, w_2, 0, 0, 0) = (w_1, w_2)$ gives $\operatorname{range} T = \mathbf{F}^2$, so $\dim \operatorname{range} T = 2$.

Problem 3B.2. Suppose $S, T \in \mathcal{L}(V)$ are such that $\operatorname{range} S \subseteq \operatorname{null} T$. Prove that $(ST)^2 = 0$.

Solution. For every $v \in V$, the vector $S(Tv)$ lies in $\operatorname{range} S \subseteq \operatorname{null} T$, so $T(S(Tv)) = 0$ and hence

$$(ST)^2 v = S\Big(T\big(S(Tv)\big)\Big) = S(0) = 0$$

by 3.10. Thus $(ST)^2 = 0$.

Problem 3B.3. Suppose $v_1, \ldots, v_m$ is a list of vectors in V. Define $T \in \mathcal{L}(\mathbf{F}^m, V)$ by

$$T(z_1, \ldots, z_m) = z_1 v_1 + \cdots + z_m v_m.$$

(a) What property of T corresponds to $v_1, \ldots, v_m$ spanning V?
(b) What property of T corresponds to the list $v_1, \ldots, v_m$ being linearly independent?

Solution. (a) Surjectivity of T. Indeed

$$\begin{aligned} \operatorname{range} T &= \{z_1 v_1 + \cdots + z_m v_m : z_k \in \mathbf{F}\} \\ &= \operatorname{span}(v_1, \ldots, v_m), \end{aligned}$$

so $\operatorname{range} T = V$, that is T surjective (3.19), says exactly that $v_1, \ldots, v_m$ spans V.
(b) Injectivity of T. Indeed, writing $z = (z_1, \ldots, z_m)$,

$$\operatorname{null} T = \{z \in \mathbf{F}^m : z_1 v_1 + \cdots + z_m v_m = 0\},$$

so $\operatorname{null} T = \{0\}$, that is T injective (3.15), says exactly that the only scalars with $z_1 v_1 + \cdots + z_m v_m = 0$ are $z_1 = \cdots = z_m = 0$, i.e. that the list is linearly independent.

Problem 3B.4. Show that $\{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$ is not a subspace of $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$.

Solution. The set $E = \{T : \dim \operatorname{null} T > 2\}$ is not closed under addition, hence is not a subspace. Take the linear maps (Check!) $S, T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$ given by

$$S(x_1, \ldots, x_5) = (x_1, x_2, 0, 0), \qquad T(x_1, \ldots, x_5) = (0, 0, x_3, x_4).$$

Then $Sx = 0$ iff $x_1 = x_2 = 0$ and $Tx = 0$ iff $x_3 = x_4 = 0$, so

$$\begin{aligned} \operatorname{null} S &= \{(0, 0, x_3, x_4, x_5)\} = \operatorname{span}(e_3, e_4, e_5), \\ \operatorname{null} T &= \{(x_1, x_2, 0, 0, x_5)\} = \operatorname{span}(e_1, e_2, e_5), \end{aligned}$$

each of dimension $3 > 2$, the spanning lists being sublists of the standard basis $e_1, \ldots, e_5$ of $\mathbf{R}^5$ (2.27); so $S, T \in E$. But $(S + T)(x_1, \ldots, x_5) = (x_1, x_2, x_3, x_4)$ vanishes only when $x_1 = x_2 = x_3 = x_4 = 0$, so $\operatorname{null}(S + T) = \operatorname{span}(e_5)$ has dimension 1 and $S + T \notin E$.

Problem 3B.5. Give an example of $T \in \mathcal{L}(\mathbf{R}^4)$ such that $\operatorname{range} T = \operatorname{null} T$.

Solution. Take the linear map (Check!) $T \in \mathcal{L}(\mathbf{R}^4)$ given by

$$T(x_1, x_2, x_3, x_4) = (0, 0, x_1, x_2).$$

Then $Tx = 0$ exactly when $x_1 = x_2 = 0$, while every output has its first two coordinates 0 and $T(a, b, 0, 0) = (0, 0, a, b)$ for all $a, b \in \mathbf{R}$; hence

$$\operatorname{null} T = \{(0, 0, x_3, x_4)\} = \{(0, 0, a, b)\} = \operatorname{range} T.$$

Problem 3B.6. Prove that there does not exist $T \in \mathcal{L}(\mathbf{R}^5)$ such that $\operatorname{range} T = \operatorname{null} T$.

Solution. If $T \in \mathcal{L}(\mathbf{R}^5)$ satisfied $\operatorname{range} T = \operatorname{null} T$, then these equal subspaces would share a common dimension k, a nonnegative integer, and the fundamental theorem of linear maps 3.21 (applicable since $\mathbf{R}^5$ is finite-dimensional) would give

$$5 = \dim \mathbf{R}^5 = \dim \operatorname{null} T + \dim \operatorname{range} T = 2k$$

which is impossible. Hence no such T exists.

Problem 3B.7. Suppose V and W are finite-dimensional with $2 \leq \dim V \leq \dim W$. Show that $\{T \in \mathcal{L}(V, W) : T \text{ is not injective}\}$ is not a subspace of $\mathcal{L}(V, W)$.

Solution. The set E of non-injective maps is not closed under addition, hence is not a subspace. Fix bases $v_1, \ldots, v_n$ of V and $w_1, \ldots, w_m$ of W, where $2 \leq n = \dim V \leq m = \dim W$, so that $w_1, \ldots, w_n$ is a linearly independent sublist; the linear map lemma 3.4 supplies $S, T \in \mathcal{L}(V, W)$ with

$$\begin{aligned} Sv_1 &= 0, & Sv_j &= w_j \quad (j = 2, \ldots, n), \\ Tv_1 &= w_1, & Tv_j &= 0 \quad (j = 2, \ldots, n). \end{aligned}$$

Here $Sv_1 = 0$ and $Tv_2 = 0$ (the vector v_2 exists because $n \geq 2$), with $v_1, v_2 \neq 0$ as members of a basis, so $\operatorname{null} S$ and $\operatorname{null} T$ are nonzero and $S, T \in E$ by 3.15. But $(S + T)v_j = w_j$ for every j, so any $v = a_1 v_1 + \cdots + a_n v_n \in \operatorname{null}(S + T)$ satisfies

$$0 = (S + T)v = a_1 w_1 + \cdots + a_n w_n,$$

forcing $a_1 = \cdots = a_n = 0$ and $v = 0$. Hence $S + T$ is injective by 3.15, and $S + T \notin E$.

Problem 3B.8. Suppose V and W are finite-dimensional with $\dim V \geq \dim W \geq 2$. Show that $\{T \in \mathcal{L}(V,W) : T \text{ is not surjective}\}$ is not a subspace of $\mathcal{L}(V,W)$.

Solution. The set E of non-surjective maps is not closed under addition, hence is not a subspace. Fix bases $v_1, \dots, v_n$ of V and $w_1, \dots, w_m$ of W, where $n = \dim V \geq m = \dim W \geq 2$, and let the linear map lemma 3.4 supply $S, T \in \mathcal{L}(V,W)$ with

$$\begin{aligned} Sv_k &= w_k \ \ (1 \leq k \leq m-1), & Sv_k &= 0 \ \ (m \leq k \leq n), \\ Tv_m &= w_m, & Tv_k &= 0 \ \ (k \neq m), \end{aligned}$$

which is legitimate because $m - 1 \geq 1$ and $m \leq n$. Expanding $v = a_1v_1 + \cdots + a_nv_n$ and using linearity gives

$$\begin{aligned} \operatorname{range} S &= \operatorname{span}(w_1, \dots, w_{m-1}), \\ \operatorname{range} T &= \operatorname{span}(w_m), \end{aligned}$$

of dimensions $m-1$ and 1, both smaller than $m = \dim W$ (the spanning lists are linearly independent, being sublists of a basis), so $\operatorname{range} S, \operatorname{range} T \neq W$ and $S, T \in E$. But $(S+T)v_k = w_k$ for $k = 1, \dots, m$, so the subspace $\operatorname{range}(S+T)$ of W (3.18) contains the basis $w_1, \dots, w_m$ and hence equals W; thus $S + T$ is surjective and $S + T \notin E$.

Problem 3B.9. Suppose $T \in \mathcal{L}(V,W)$ is injective and $v_1, \dots, v_n$ is linearly independent in V. Prove that $Tv_1, \dots, Tv_n$ is linearly independent in W.

Solution. Suppose $a_1, \dots, a_n \in \mathbf{F}$ satisfy $a_1Tv_1 + \cdots + a_nTv_n = 0$. By linearity of T this says $a_1v_1 + \cdots + a_nv_n \in \operatorname{null} T$, and $\operatorname{null} T = \{0\}$ by 3.15 since T is injective, so

$$a_1v_1 + \cdots + a_nv_n = 0.$$

Linear independence of $v_1, \dots, v_n$ now forces $a_1 = \cdots = a_n = 0$. Hence $Tv_1, \dots, Tv_n$ is linearly independent in W.

Problem 3B.10. Suppose $v_1, \dots, v_n$ spans V and $T \in \mathcal{L}(V,W)$. Show that $Tv_1, \dots, Tv_n$ spans $\operatorname{range} T$.

Solution. Each Tv_k lies in $\operatorname{range} T$, which is a subspace of W by 3.18 and therefore contains every linear combination of them; hence $\operatorname{span}(Tv_1, \dots, Tv_n) \subseteq \operatorname{range} T$. Conversely, every $w \in \operatorname{range} T$ is $w = Tv$ with $v = a_1v_1 + \cdots + a_nv_n$ for some $a_k \in \mathbf{F}$, the list $v_1, \dots, v_n$ spanning V, so linearity gives

$$w = T(a_1v_1 + \cdots + a_nv_n) = a_1Tv_1 + \cdots + a_nTv_n$$

in $\operatorname{span}(Tv_1, \dots, Tv_n)$. The two inclusions give $\operatorname{span}(Tv_1, \dots, Tv_n) = \operatorname{range} T$.

Problem 3B.11. Suppose that V is finite-dimensional and that $T \in \mathcal{L}(V,W)$. Prove that there exists a subspace U of V such that

$$U \cap \operatorname{null} T = \{0\} \quad \text{and} \quad \operatorname{range} T = \{Tu : u \in U\}.$$

Solution. Take $U = \operatorname{span}(w_1, \dots, w_n)$, where $u_1, \dots, u_m$ is a basis of $\operatorname{null} T$ (finite-dimensional by 2.25, being a subspace of the finite-dimensional V), extended by 2.32 to a basis

$$u_1, \dots, u_m, w_1, \dots, w_n$$

of V.

If $v \in U \cap \operatorname{null} T$, writing v in each of the two bases and subtracting gives

$$a_1 u_1 + \cdots + a_m u_m - b_1 w_1 - \cdots - b_n w_n = 0,$$

so every coefficient vanishes by independence of the extended list; hence $U \cap \operatorname{null} T = \{0\}$.
For the range, $\{Tu : u \in U\} \subseteq \operatorname{range} T$ is immediate, and given $x = Tv$ with $v = a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_n$, the vector $u = b_1 w_1 + \cdots + b_n w_n \in U$ satisfies

$$Tu = Tv - (a_1 T u_1 + \cdots + a_m T u_m) = Tv = x$$

since each $u_j \in \operatorname{null} T$. Hence $\operatorname{range} T = \{Tu : u \in U\}$.

Problem 3B.12. Suppose T is a linear map from $\mathbf{F}^4$ to $\mathbf{F}^2$ such that

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\}.$$

Prove that T is surjective.

Solution. Here $\dim \operatorname{null} T = 2$: a vector lies in $\operatorname{null} T$ exactly when it has the form

$$(5x_2, x_2, 7x_4, x_4) = x_2(5, 1, 0, 0) + x_4(0, 0, 7, 1),$$

and $(5, 1, 0, 0), (0, 0, 7, 1)$ is linearly independent (read off the second and fourth coordinates), hence a basis of $\operatorname{null} T$. Since $\dim \mathbf{F}^4 = 4$, the fundamental theorem of linear maps (3.21) gives

$$\dim \operatorname{range} T = 4 - 2 = 2 = \dim \mathbf{F}^2.$$

As $\operatorname{range} T$ is a subspace of $\mathbf{F}^2$ (by 3.18) of full dimension, 2.39 gives $\operatorname{range} T = \mathbf{F}^2$; that is, T is surjective.

Problem 3B.13. Suppose U is a three-dimensional subspace of $\mathbf{R}^8$ and that T is a linear map from $\mathbf{R}^8$ to $\mathbf{R}^5$ such that $\operatorname{null} T = U$. Prove that T is surjective.

Solution. The fundamental theorem of linear maps (3.21), applied to T on the finite-dimensional space $\mathbf{R}^8$, gives

$$\dim \operatorname{range} T = 8 - \dim \operatorname{null} T = 8 - \dim U = 5.$$

Since $\operatorname{range} T$ is a subspace of $\mathbf{R}^5$ (by 3.18) with $\dim \operatorname{range} T = 5 = \dim \mathbf{R}^5$, 2.39 gives $\operatorname{range} T = \mathbf{R}^5$. Hence T is surjective.

Problem 3B.14. Prove that there does not exist a linear map from $\mathbf{F}^5$ to $\mathbf{F}^2$ whose null space equals $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\}$.

Solution. Such a map would need a three-dimensional range inside $\mathbf{F}^2$. Write N for the displayed set; a vector lies in N exactly when it has the form

$$(3x_2, x_2, x_3, x_3, x_3) = x_2(3, 1, 0, 0, 0) + x_3(0, 0, 1, 1, 1),$$

and $(3, 1, 0, 0, 0), (0, 0, 1, 1, 1)$ is linearly independent (read off the second and third coordinates), so it is a basis of N and $\dim N = 2$.
If $T \in \mathcal{L}(\mathbf{F}^5, \mathbf{F}^2)$ had $\operatorname{null} T = N$, then 3.21 applied to T on the finite-dimensional $\mathbf{F}^5$ would give

$$\dim \operatorname{range} T = 5 - 2 = 3,$$

contradicting $\dim \operatorname{range} T \le \dim \mathbf{F}^2 = 2$ (by 2.37, since $\operatorname{range} T$ is a subspace of $\mathbf{F}^2$ by 3.18).

Problem 3B.15. Suppose there exists a linear map on V whose null space and range are both finite-dimensional. Prove that V is finite-dimensional.

Solution. The list $u_1, \ldots, u_m, v_1, \ldots, v_n$ spans V, where $T \in \mathcal{L}(V)$ is the given map, $u_1, \ldots, u_m$ is a basis of $\operatorname{null} T$ and $w_1, \ldots, w_n$ a basis of $\operatorname{range} T$ (both exist by 2.31, these spaces being finite-dimensional by hypothesis), and $v_k \in V$ is chosen with $Tv_k = w_k$.
Indeed, given $v \in V$, write $Tv = b_1 w_1 + \cdots + b_n w_n$; then linearity gives

$$T\big(v - (b_1 v_1 + \cdots + b_n v_n)\big) = Tv - (b_1 w_1 + \cdots + b_n w_n) = 0,$$

so $v - (b_1 v_1 + \cdots + b_n v_n) = a_1 u_1 + \cdots + a_m u_m$ for some scalars a_j, whence

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_n v_n.$$

Thus V is spanned by a finite list, so V is finite-dimensional.

Problem 3B.16. Suppose V and W are both finite-dimensional. Prove that there exists an injective linear map from V to W if and only if $\dim V \leq \dim W$.

Solution. (i) If $T \in \mathcal{L}(V, W)$ is injective, then $\operatorname{null} T = \{0\}$ (by 3.15), so 3.21 and 2.37 give

$$\dim V = \dim \operatorname{range} T \leq \dim W,$$

the inequality because $\operatorname{range} T$ is a subspace of the finite-dimensional W (by 3.18).
(ii) Conversely, let $n = \dim V \leq m = \dim W$, let $v_1, \ldots, v_n$ be a basis of V and $w_1, \ldots, w_m$ a basis of W, and use the linear map lemma (3.4) to get $T \in \mathcal{L}(V, W)$ with $Tv_k = w_k$ for $k = 1, \ldots, n$ (legitimate since $n \leq m$). If $v = c_1 v_1 + \cdots + c_n v_n \in \operatorname{null} T$, then

$$0 = Tv = c_1 w_1 + \cdots + c_n w_n,$$

so all $c_k = 0$ by linear independence of the sublist $w_1, \ldots, w_n$; hence $v = 0$ and T is injective by 3.15.

Problem 3B.17. Suppose V and W are both finite-dimensional. Prove that there exists a surjective linear map from V onto W if and only if $\dim V \geq \dim W$.

Solution. (i) If $T \in \mathcal{L}(V, W)$ is surjective, then $\operatorname{range} T = W$, so 3.21 gives

$$\dim V = \dim \operatorname{null} T + \dim W \geq \dim W.$$

(ii) Conversely, let $n = \dim V \geq m = \dim W$, let $v_1, \ldots, v_n$ be a basis of V and $w_1, \ldots, w_m$ a basis of W, and use the linear map lemma (3.4) to get $T \in \mathcal{L}(V, W)$ with

$$Tv_k = \begin{cases} w_k & \text{if } 1 \leq k \leq m, \\ 0 & \text{if } m < k \leq n \end{cases}$$

(legitimate since $m \leq n$). Then $\operatorname{range} T$ is a subspace of W (by 3.18) containing $w_1, \ldots, w_m$, hence containing $\operatorname{span}(w_1, \ldots, w_m) = W$ by 2.6. So T is surjective.

Problem 3B.18. Suppose V and W are finite-dimensional and that U is a subspace of V. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$ if and only if $\dim U \geq \dim V - \dim W$.

Solution. (i) If $T \in \mathcal{L}(V, W)$ has $\operatorname{null} T = U$, then 3.21 and 2.37 (applied to the subspace $\operatorname{range} T$ of the finite-dimensional W, by 3.18) give

$$\dim V = \dim U + \dim \operatorname{range} T \leq \dim U + \dim W,$$

which rearranges to $\dim U \geq \dim V - \dim W$.
(ii) Conversely, suppose $\dim U \geq \dim V - \dim W$. Extend a basis $u_1, \ldots, u_j$ of U (finite-dimensional by 2.25) to a basis $u_1, \ldots, u_j, v_1, \ldots, v_n$ of V (by 2.32); then $n = \dim V - \dim U \leq \dim W$, so a basis $w_1, \ldots, w_m$ of W has the linearly independent sublist $w_1, \ldots, w_n$. By 3.4 there is $T \in \mathcal{L}(V, W)$ with

$$Tu_i = 0 \quad (i = 1, \ldots, j), \qquad Tv_k = w_k \quad (k = 1, \ldots, n).$$

Then $U = \operatorname{span}(u_1, \ldots, u_j) \subseteq \operatorname{null} T$, the latter being a subspace by 3.13; and if $v = a_1u_1 + \cdots + a_ju_j + b_1v_1 + \cdots + b_nv_n$ lies in $\operatorname{null} T$, then

$$0 = Tv = b_1w_1 + \cdots + b_nw_n,$$

forcing every $b_k = 0$, so $v \in U$. Hence $\operatorname{null} T = U$.

Problem 3B.19. Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is injective if and only if there exists $S \in \mathcal{L}(W, V)$ such that ST is the identity operator on V.

Solution. (i) If $ST = I$ and $Tu = Tv$, then applying S gives

$$u = (ST)u = S(Tu) = S(Tv) = (ST)v = v,$$

so T is injective.
(ii) Conversely, suppose T is injective. Extend a basis $w_1, \ldots, w_n$ of $\operatorname{range} T$ (a subspace of the finite-dimensional W, by 3.18 and 2.25) to a basis $w_1, \ldots, w_m$ of W (by 2.32), choose $v_k \in V$ with $Tv_k = w_k$ for $k \leq n$, and use the linear map lemma (3.4) to get $S \in \mathcal{L}(W, V)$ with

$$Sw_k = \begin{cases} v_k & \text{if } 1 \leq k \leq n, \\ 0 & \text{if } n < k \leq m. \end{cases}$$

Given $v \in V$, write $Tv = c_1w_1 + \cdots + c_nw_n$. Then

$$(ST)v = c_1v_1 + \cdots + c_nv_n, \qquad T(c_1v_1 + \cdots + c_nv_n) = Tv,$$

so injectivity of T gives $c_1v_1 + \cdots + c_nv_n = v$ and hence $(ST)v = v$. Thus $ST = I$.

Problem 3B.20. Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is surjective if and only if there exists $S \in \mathcal{L}(W, V)$ such that TS is the identity operator on W.

Solution. (i) If $TS = I$, then every $w \in W$ satisfies

$$w = (TS)w = T(Sw) \in \operatorname{range} T,$$

so $\operatorname{range} T = W$; that is, T is surjective.
(ii) Conversely, suppose T is surjective. Let $w_1, \ldots, w_m$ be a basis of W (by 2.31, W being finite-dimensional), choose $v_k \in V$ with $Tv_k = w_k$ (possible since $\operatorname{range} T = W$), and use the linear map lemma (3.4) to get $S \in \mathcal{L}(W, V)$ with $Sw_k = v_k$ for $k = 1, \ldots, m$. Then

$$(TS)w_k = T(Sw_k) = Tv_k = w_k$$

for every k, so TS and I are linear maps on W agreeing on a basis of W; by the uniqueness assertion of 3.4 they are equal, so $TS = I$.

Problem 3B.21. Suppose V is finite-dimensional, $T \in \mathcal{L}(V, W)$, and U is a subspace of W. Prove that $\{v \in V : Tv \in U\}$ is a subspace of V and

$$\dim\{v \in V : Tv \in U\} = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

Solution. Write $E = \{v \in V : Tv \in U\}$. Then E is a subspace of V by 1.34: $T0 = 0 \in U$ (by 3.10, and since U is a subspace of W), while for $v_1, v_2, v \in E$ and $\lambda \in \mathbf{F}$,

$$T(v_1 + v_2) = Tv_1 + Tv_2 \in U, \qquad T(\lambda v) = \lambda Tv \in U,$$

U being closed under addition and scalar multiplication.
Now apply the fundamental theorem of linear maps (3.21) to the restriction $T|_E \in \mathcal{L}(E, W)$, whose domain is finite-dimensional by 2.25:

$$\dim E = \dim \operatorname{null}(T|_E) + \dim \operatorname{range}(T|_E).$$

Here $\operatorname{null}(T|_E) = \operatorname{null} T$, since $Tv = 0$ forces $Tv = 0 \in U$ and hence $v \in E$; and $\operatorname{range}(T|_E) = U \cap \operatorname{range} T$, since $Tv \in U$ for every $v \in E$, while any $w \in U \cap \operatorname{range} T$ equals Tv for some v, and that v lies in E because $Tv = w \in U$. Substituting,

$$\dim E = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T),$$

which is the asserted formula.

Problem 3B.22. Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T.$$

Solution. Apply the fundamental theorem of linear maps (3.21) to the restriction

$$R = T|_{\operatorname{null} ST} \in \mathcal{L}(\operatorname{null} ST,\ V),$$

whose domain is finite-dimensional (a subspace of the finite-dimensional U, by 3.13 and 2.25):

$$\dim \operatorname{null} ST = \dim \operatorname{null} R + \dim \operatorname{range} R.$$

Now $\operatorname{null} R \subseteq \operatorname{null} T$, since $Ru = Tu$; and $\operatorname{range} R \subseteq \operatorname{null} S$, since $v = Tu$ with $u \in \operatorname{null} ST$ gives

$$Sv = S(Tu) = (ST)u = 0.$$

Each containment sits inside a finite-dimensional space ($\operatorname{null} T$ inside U, and $\operatorname{null} S$ inside V, by 3.13 and 2.25), so 2.37 gives $\dim \operatorname{null} R \leq \dim \operatorname{null} T$ and $\dim \operatorname{range} R \leq \dim \operatorname{null} S$. Substituting,

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T.$$

Problem 3B.23. Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{range} ST \leq \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

Solution. (i) $\operatorname{range} ST \subseteq \operatorname{range} S$, since $(ST)u = S(Tu)$ with $Tu \in V$. Both are subspaces of W (by 3.18) and $\operatorname{range} S$ is finite-dimensional (by 3.21 applied to S on the finite-dimensional V), so 2.37 gives $\dim \operatorname{range} ST \leq \dim \operatorname{range} S$.
(ii) Restrict S to the subspace $\operatorname{range} T$ of V (a subspace by 3.18), obtaining $R = S|_{\operatorname{range} T}$, whose range is

$$\operatorname{range} R = \{S(Tu) : u \in U\} = \operatorname{range} ST,$$

since v ranges over $\operatorname{range} T$ exactly when $v = Tu$ for some $u \in U$. As $\operatorname{range} T$ is finite-dimensional (by 3.21 applied to T on the finite-dimensional U), 3.21 applied to R gives

$$\begin{aligned}\dim \operatorname{range} T &= \dim \operatorname{null} R + \dim \operatorname{range} ST \\ &\geq \dim \operatorname{range} ST.\end{aligned}$$

Being at most each of the two dimensions, $\dim \operatorname{range} ST$ is at most their minimum.

Problem 3B.24. (a) Suppose $\dim V = 5$ and $S, T \in \mathcal{L}(V)$ are such that $ST = 0$. Prove that $\dim \operatorname{range} TS \leq 2$.
(b) Give an example of $S, T \in \mathcal{L}(\mathbf{F}^5)$ with $ST = 0$ and $\dim \operatorname{range} TS = 2$.

Solution. (a) The hypothesis $ST = 0$ says $\operatorname{range} T \subseteq \operatorname{null} S$, since $S(Tv) = (ST)v = 0$ for every $v \in V$. Both are subspaces of the five-dimensional V (by 3.18 and 3.13), so 2.37 and the fundamental theorem of linear maps (3.21) applied to S give

$$\begin{aligned}\dim \operatorname{range} T &\leq \dim \operatorname{null} S \\ &= 5 - \dim \operatorname{range} S,\end{aligned}$$

that is, $\dim \operatorname{range} S + \dim \operatorname{range} T \leq 5$. Hence the two ranges cannot both have dimension at least 3, so Exercise 3B.23 applied to TS (legitimate, as V is finite-dimensional) gives

$$\begin{aligned}\dim \operatorname{range} TS &\leq \min\{\dim \operatorname{range} T, \dim \operatorname{range} S\} \\ &\leq 2.\end{aligned}$$

(b) Take $S, T \in \mathcal{L}(\mathbf{F}^5)$ given by

$$\begin{aligned}S(x_1, \ldots, x_5) &= (x_1, x_2, 0, 0, 0), \\ T(x_1, \ldots, x_5) &= (0, 0, x_1, x_2, 0),\end{aligned}$$

each linear because every output coordinate is a linear function of the input. Then

$$(ST)(x_1, \ldots, x_5) = S(0, 0, x_1, x_2, 0) = 0,$$

so $ST = 0$, while $(TS)(x_1, \ldots, x_5) = T(x_1, x_2, 0, 0, 0) = (0, 0, x_1, x_2, 0)$, so

$$\operatorname{range} TS = \operatorname{span}\big((0, 0, 1, 0, 0),\ (0, 0, 0, 1, 0)\big)$$

has dimension 2, those two vectors being linearly independent.

Problem 3B.25. Suppose that W is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} S \subseteq \operatorname{null} T$ if and only if there exists $E \in \mathcal{L}(W)$ such that $T = ES$.

Solution. (i) If $T = ES$ and $v \in \operatorname{null} S$, then

$$Tv = (ES)v = E(Sv) = E0 = 0$$

by 3.10, so $\operatorname{null} S \subseteq \operatorname{null} T$.
(ii) Conversely, suppose $\operatorname{null} S \subseteq \operatorname{null} T$. Let $w_1, \ldots, w_m$ be a basis of $\operatorname{range} S$ (a subspace of W by 3.18, hence finite-dimensional by 2.25; the empty list if $\operatorname{range} S = \{0\}$), choose $v_k \in V$ with $Sv_k = w_k$, extend by 2.32 to a basis $w_1, \ldots, w_n$ of W, and use the linear map lemma (3.4) to get $E \in \mathcal{L}(W)$ with

$$Ew_k = Tv_k \quad (k \leq m), \qquad Ew_k = 0 \quad (m < k \leq n).$$

Given $v \in V$, write $Sv = a_1 w_1 + \cdots + a_m w_m$ and put $u = a_1 v_1 + \cdots + a_m v_m$. Then $Su = a_1 w_1 + \cdots +$

$a_m w_m = Sv$, so $v - u \in \operatorname{null} S \subseteq \operatorname{null} T$ and hence $Tv = Tu$. Therefore

$$(ES)v = a_1 Ew_1 + \cdots + a_m Ew_m = a_1 Tv_1 + \cdots + a_m Tv_m = Tu = Tv,$$

so $ES = T$.

Problem 3B.26. Suppose that V is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} S \subseteq \operatorname{range} T$ if and only if there exists $E \in \mathcal{L}(V)$ such that $S = TE$.

Solution. (i) If $S = TE$, then every $w = Sv$ in $\operatorname{range} S$ satisfies

$$w = (TE)v = T(Ev) \in \operatorname{range} T,$$

so $\operatorname{range} S \subseteq \operatorname{range} T$.
(ii) Conversely, suppose $\operatorname{range} S \subseteq \operatorname{range} T$. Let $v_1, \ldots, v_n$ be a basis of V (which exists by 2.31, V being finite-dimensional), choose $u_k \in V$ with $Tu_k = Sv_k$ (possible because $Sv_k \in \operatorname{range} S \subseteq \operatorname{range} T$), and use the linear map lemma (3.4) to get $E \in \mathcal{L}(V)$ with $Ev_k = u_k$ for $k = 1, \ldots, n$. Then

$$(TE)v_k = T(Ev_k) = Tu_k = Sv_k$$

for every k, so TE and S are linear maps from V to W agreeing on a basis of V; by the uniqueness assertion of 3.4 they are equal, so $S = TE$.

Problem 3B.27. Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \operatorname{null} P \oplus \operatorname{range} P$.

Solution. Every $v \in V$ splits as $v = (v - Pv) + Pv$, with $Pv \in \operatorname{range} P$ and

$$P(v - Pv) = Pv - P(Pv) = Pv - P^2 v = 0,$$

so $v - Pv \in \operatorname{null} P$; hence $V = \operatorname{null} P + \operatorname{range} P$, both summands being subspaces of V (by 3.13 and 3.18).
If $v \in \operatorname{null} P \cap \operatorname{range} P$, say $v = Pu$, then

$$v = Pu = P^2 u = P(Pu) = Pv = 0,$$

so $\operatorname{null} P \cap \operatorname{range} P = \{0\}$. By 1.46 the sum is therefore direct:

$$V = \operatorname{null} P \oplus \operatorname{range} P.$$

Problem 3B.28. Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is such that $\deg Dp = (\deg p) - 1$ for every nonconstant polynomial $p \in \mathcal{P}(\mathbf{R})$. Prove that D is surjective.
The notation D is used above to remind you of the differentiation map that sends a polynomial p to p'.

Solution. It suffices to show that $D_m = D|_{\mathcal{P}_m(\mathbf{R})}$ maps onto $\mathcal{P}_{m-1}(\mathbf{R})$ for every $m \geq 1$, since every $q \in \mathcal{P}(\mathbf{R})$ lies in $\mathcal{P}_{m-1}(\mathbf{R})$ for some $m \geq 1$ (take $m = \deg q + 1$, or $m = 1$ if $q = 0$). Throughout, $\dim \mathcal{P}_m(\mathbf{R}) = m + 1$ (by 2.36), $\deg 0 = -\infty$, and nonconstant means $\deg p \geq 1$.
First, $D1$ is constant: both x and $x + 1$ are nonconstant of degree 1, so $\deg Dx = \deg D(x + 1) = 0$, and linearity gives

$$D1 = D(x + 1) - Dx,$$

a difference of two constants. Hence D maps $\mathcal{P}_m(\mathbf{R})$ into $\mathcal{P}_{m-1}(\mathbf{R})$ for each $m \geq 1$: a nonconstant $p \in \mathcal{P}_m(\mathbf{R})$ has $\deg Dp = \deg p - 1 \leq m - 1$, while a constant p has Dp constant, so $Dp \in \mathcal{P}_0(\mathbf{R}) \subseteq \mathcal{P}_{m-1}(\mathbf{R})$.
Also $\operatorname{null} D \subseteq \mathcal{P}_0(\mathbf{R})$, since a nonconstant p has $\deg Dp = \deg p - 1 \geq 0$ and hence $Dp \neq 0$. So

$D_m \in \mathcal{L}\big(\mathcal{P}_m(\mathbf{R}),\ \mathcal{P}_{m-1}(\mathbf{R})\big)$ has $\dim \operatorname{null} D_m \le \dim \mathcal{P}_0(\mathbf{R}) = 1$ (by 2.37), and 3.21 applied to D_m on the finite-dimensional $\mathcal{P}_m(\mathbf{R})$ gives

$$\dim \operatorname{range} D_m \ge (m+1) - 1 = m = \dim \mathcal{P}_{m-1}(\mathbf{R}).$$

Since $\operatorname{range} D_m$ is a subspace of $\mathcal{P}_{m-1}(\mathbf{R})$ (by 3.18), 2.37 forces equality here, so 2.39 gives $\operatorname{range} D_m = \mathcal{P}_{m-1}(\mathbf{R})$, and D is surjective.

Problem 3B.29. Suppose $p \in \mathcal{P}(\mathbb{R})$. Prove that there exists a polynomial $q \in \mathcal{P}(\mathbb{R})$ such that $5q'' + 3q' = p$.
This exercise can be done without linear algebra, but it's more fun to do it using linear algebra.

Solution. Fix $m \ge 0$ with $p \in \mathcal{P}_m(\mathbb{R})$ and apply the fundamental theorem of linear maps to $T \in \mathcal{L}\big(\mathcal{P}_{m+1}(\mathbb{R}),\ \mathcal{P}_m(\mathbb{R})\big)$ given by

$$Tq = 5q'' + 3q',$$

which indeed lands in $\mathcal{P}_m(\mathbb{R})$ (differentiation lowers degree by at least one) and is linear (differentiation is linear).
Here $\operatorname{null} T = \mathcal{P}_0(\mathbb{R})$, of dimension 1: every constant polynomial lies in it, and conversely $5q'' + 3q' = 0$ makes $r = q'$ satisfy $r = -\frac{5}{3}r'$, which is impossible for $r \neq 0$ since then $\deg r' < \deg r$; so $q' = 0$. Hence 3.21 gives

$$\dim \operatorname{range} T = (m+2) - 1 = m+1 = \dim \mathcal{P}_m(\mathbb{R}),$$

using $\dim \mathcal{P}_k(\mathbb{R}) = k+1$ (by 2.36). As $\operatorname{range} T$ is a subspace of $\mathcal{P}_m(\mathbb{R})$ (by 3.18) of full dimension, 2.39 gives $\operatorname{range} T = \mathcal{P}_m(\mathbb{R})$, which contains p; so some $q \in \mathcal{P}_{m+1}(\mathbb{R})$ has $5q'' + 3q' = p$.

Problem 3B.30. Suppose $\varphi \in \mathcal{L}(V, \mathbb{F})$ and $\varphi \neq 0$. Suppose $u \in V$ is not in $\operatorname{null} \varphi$. Prove that

$$V = \operatorname{null} \varphi \oplus \{au : a \in \mathbb{F}\}.$$

Solution. Write $U = \{au : a \in \mathbb{F}\} = \operatorname{span}(u)$, a subspace of V by 2.6, and note that the hypothesis $u \notin \operatorname{null} \varphi$ says $\varphi(u) \neq 0$.
Given $v \in V$, put $a = \varphi(v)/\varphi(u) \in \mathbb{F}$ (legitimate since $\varphi(u) \neq 0$); linearity of φ gives

$$\varphi(v - au) = \varphi(v) - a\,\varphi(u) = 0,$$

so $v = (v - au) + au \in \operatorname{null} \varphi + U$, whence $V = \operatorname{null} \varphi + U$.
If $w \in (\operatorname{null} \varphi) \cap U$, say $w = au$, then

$$0 = \varphi(w) = a\,\varphi(u),$$

so $a = 0$ and $w = 0$; thus $(\operatorname{null} \varphi) \cap U = \{0\}$. By 1.46, applied to the subspaces $\operatorname{null} \varphi$ (a subspace by 3.13) and U, the sum is direct:

$$V = \operatorname{null} \varphi \oplus \{au : a \in \mathbb{F}\}.$$

Problem 3B.31. Suppose V is finite-dimensional, X is a subspace of V, and Y is a finite-dimensional subspace of W. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = X$ and $\operatorname{range} T = Y$ if and only if $\dim X + \dim Y = \dim V$.

Solution. (i) If $T \in \mathcal{L}(V, W)$ has $\operatorname{null} T = X$ and $\operatorname{range} T = Y$, then the fundamental theorem of linear maps (3.21), applicable since V is finite-dimensional, gives

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T = \dim X + \dim Y.$$

(ii) Conversely, suppose $\dim X + \dim Y = \dim V$; write $n = \dim V$ and $m = \dim X$ (finite by 2.25), so

$\dim Y = n - m$. Extend a basis $x_1, \ldots, x_m$ of X to a basis

$$x_1, \ldots, x_m, v_1, \ldots, v_{n-m}$$

of V (by 2.32, with exactly $n - m$ vectors added since every basis of V has length n, by 2.34), take a basis $y_1, \ldots, y_{n-m}$ of $Y \subseteq W$, and use the linear map lemma (3.4) to get $T \in \mathcal{L}(V, W)$ with

$$Tx_k = 0 \quad (k \le m), \qquad Tv_j = y_j \quad (j \le n - m).$$

Then $Tv = \sum_j b_j y_j$ for $v = \sum_k a_k x_k + \sum_j b_j v_j$, and each $y_j = Tv_j$, so

$$\operatorname{range} T = \operatorname{span}(y_1, \ldots, y_{n-m}) = Y.$$

Also $X = \operatorname{span}(x_1, \ldots, x_m) \subseteq \operatorname{null} T$, the latter being a subspace by 3.13, while 3.21 gives

$$\dim \operatorname{null} T = n - (n - m) = m = \dim X;$$

hence $X = \operatorname{null} T$ by 2.39.

Problem 3B.32. Suppose V is finite-dimensional with $\dim V > 1$. Show that if $\varphi \colon \mathcal{L}(V) \to \mathbb{F}$ is a linear map such that $\varphi(ST) = \varphi(S)\varphi(T)$ for all $S, T \in \mathcal{L}(V)$, then $\varphi = 0$.
Hint: The description of the two-sided ideals of $\mathcal{L}(V)$ given by Exercise 17 in Section 3A might be useful.

Solution. The set $\operatorname{null} \varphi$ is a two-sided ideal of $\mathcal{L}(V)$: it is a subspace of $\mathcal{L}(V)$ (by 3.13), and for $E \in \operatorname{null} \varphi$ and $T \in \mathcal{L}(V)$ the multiplicativity of φ gives

$$\varphi(TE) = \varphi(T)\varphi(E) = 0 = \varphi(E)\varphi(T) = \varphi(ET),$$

so $TE, ET \in \operatorname{null} \varphi$. Since V is finite-dimensional, Exercise 3A.17 says the only two-sided ideals of $\mathcal{L}(V)$ are $\{0\}$ and $\mathcal{L}(V)$.
The alternative $\operatorname{null} \varphi = \{0\}$ is impossible: it would make φ injective (by 3.15), whereas $\dim V > 1$ and 3.72 (with $W = V$) give

$$\dim \mathcal{L}(V) = (\dim V)^2 \ge 4 > 1 = \dim \mathbb{F},$$

so no linear map from $\mathcal{L}(V)$ to $\mathbb{F}$ is injective (by 3.22).
Hence $\operatorname{null} \varphi = \mathcal{L}(V)$, that is, $\varphi = 0$.

Problem 3B.33. Suppose that V and W are real vector spaces and $T \in \mathcal{L}(V, W)$. Define $T_{\mathbb{C}} \colon V_{\mathbb{C}} \to W_{\mathbb{C}}$ by

$$T_{\mathbb{C}}(u + iv) = Tu + iTv$$

for all $u, v \in V$.
(a) Show that $T_{\mathbb{C}}$ is a (complex) linear map from $V_{\mathbb{C}}$ to $W_{\mathbb{C}}$.
(b) Show that $T_{\mathbb{C}}$ is injective if and only if T is injective.
(c) Show that $\operatorname{range} T_{\mathbb{C}} = W_{\mathbb{C}}$ if and only if $\operatorname{range} T = W$.
See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbb{C}}$. The linear map $T_{\mathbb{C}}$ is called the complexification of the linear map T.

Solution. Throughout, each element of $V_{\mathbb{C}}$ is written $u + iv$ with $u, v \in V$ uniquely determined, and by Exercise 1B.8 the operations are

$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2),$$

$$(a + bi)(u + iv) = (au - bv) + i(av + bu) \qquad (a, b \in \mathbb{R}),$$

and likewise in $W_{\mathbb{C}}$; in particular $w_1 + iw_2 = 0$ exactly when $w_1 = w_2 = 0$.
(a) Additivity and complex homogeneity:

$$\begin{aligned} T_{\mathbb{C}}\big((u_1+iv_1)+(u_2+iv_2)\big) &= T(u_1+u_2) + i\,T(v_1+v_2) \\ &= (Tu_1 + iTv_1) + (Tu_2 + iTv_2) \\ &= T_{\mathbb{C}}(u_1+iv_1) + T_{\mathbb{C}}(u_2+iv_2), \end{aligned}$$

$$\begin{aligned} T_{\mathbb{C}}\big((a+bi)(u+iv)\big) &= T(au-bv) + i\,T(av+bu) \\ &= (a\,Tu - b\,Tv) + i(a\,Tv + b\,Tu) \\ &= (a+bi)\,T_{\mathbb{C}}(u+iv), \end{aligned}$$

each middle equality by $\mathbb{R}$-linearity of T and each last one by the definition of the operations in $W_{\mathbb{C}}$. Hence $T_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}}, W_{\mathbb{C}})$.
(b) Here $\operatorname{null} T_{\mathbb{C}} = (\operatorname{null} T)_{\mathbb{C}}$, since $Tu + iTv = 0$ exactly when $Tu = 0$ and $Tv = 0$. So if $\operatorname{null} T = \{0\}$, then $u + iv \in \operatorname{null} T_{\mathbb{C}}$ forces $u = v = 0$; conversely, if $\operatorname{null} T_{\mathbb{C}} = \{0\}$, then $u \in \operatorname{null} T$ gives $T_{\mathbb{C}}(u+i0) = Tu + i\,T0 = 0$, so $u + i0 = 0$ and $u = 0$. Now 3.15 converts both statements into injectivity.
(c) If $\operatorname{range} T = W$, then given $w_1 + iw_2 \in W_{\mathbb{C}}$ choose $u, v \in V$ with $Tu = w_1$ and $Tv = w_2$, so that

$$T_{\mathbb{C}}(u+iv) = Tu + iTv = w_1 + iw_2,$$

giving $\operatorname{range} T_{\mathbb{C}} = W_{\mathbb{C}}$. Conversely, if $\operatorname{range} T_{\mathbb{C}} = W_{\mathbb{C}}$, then each $w \in W$ satisfies $w + i0 = T_{\mathbb{C}}(u+iv) = Tu + iTv$ for some $u, v \in V$; comparing coordinates gives $Tu = w$, so $\operatorname{range} T = W$.

3.3 Exercises 3C

Problem 3C.1. Suppose $T \in \mathcal{L}(V, W)$. Show that with respect to each choice of bases of V and W, the matrix of T has at least $\dim \operatorname{range} T$ nonzero entries.

Solution. Fix bases $v_1, \dots, v_n$ of V and $w_1, \dots, w_m$ of W, write $A = \mathcal{M}(T)$, so that $Tv_k = \sum_{j=1}^m A_{j,k} w_j$ for each k, and let K be the set of indices k whose column of A has at least one nonzero entry. If $k \notin K$ then $Tv_k = 0$, so

$$\begin{aligned} \operatorname{range} T &= \operatorname{span}(Tv_1, \dots, Tv_n) \\ &= \operatorname{span}(Tv_k : k \in K), \end{aligned}$$

the first equality because $v_1, \dots, v_n$ spans V and T is linear. Reducing this spanning list of length $|K|$ to a basis (2.30) gives $\dim \operatorname{range} T \leq |K|$.
Finally, picking one nonzero entry from each column indexed by K exhibits $|K|$ distinct nonzero entries of A (distinct columns give distinct entries), so

$$\begin{aligned} \dim \operatorname{range} T &\leq |K| \\ &\leq \#\{\text{nonzero entries of } \mathcal{M}(T)\}. \end{aligned}$$

Problem 3C.2. Suppose $T \in \mathcal{L}(V, W)$, where V and W are finite-dimensional and nonzero. Prove that $\dim \operatorname{range} T = 1$ if and only if there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ equal 1.

Solution. Write $n = \dim V \geq 1$ and $m = \dim W \geq 1$.
(i) If every entry of $\mathcal{M}(T)$ equals 1 with respect to bases $v_1, \dots, v_n$ of V and $w_1, \dots, w_m$ of W, then $Tv_k = w_1 + \cdots + w_m = w$ for every k, and $w \neq 0$ since $w_1, \dots, w_m$ is linearly independent. As $v_1, \dots, v_n$ spans V,

$$\begin{aligned} \operatorname{range} T &= \operatorname{span}(Tv_1, \dots, Tv_n) \\ &= \operatorname{span}(w), \end{aligned}$$

which has dimension 1.

(ii) Conversely, suppose $\dim \operatorname{range} T = 1$, say $\operatorname{range} T = \operatorname{span}(w)$ with $w \neq 0$. Extend to a basis $u_1, \dots, u_m$ of W with $u_1 = w$ (by 2.32) and set

$$w_1 = u_1 - u_2 - \cdots - u_m, \qquad w_j = u_j \quad (2 \le j \le m).$$

Then $\operatorname{span}(w_1, \dots, w_m)$ contains $u_2, \dots, u_m$ and $w_1 + u_2 + \cdots + u_m = u_1$, so it equals W; having length $m = \dim W$, this list is a basis of W (by 2.42), and $w_1 + \cdots + w_m = u_1 = w$.
For V, choose v_1 with $Tv_1 = w$ and a basis $z_1, \dots, z_{n-1}$ of $\operatorname{null} T$, which has dimension $n-1$ by 3.21. The list $v_1, z_1, \dots, z_{n-1}$ is linearly independent (applying T to a vanishing combination gives $aw = 0$, so $a = 0$, and then the remaining coefficients vanish), hence a basis of V by 2.38. Set

$$v_k = v_1 + z_{k-1} \qquad (2 \le k \le n).$$

Then $\operatorname{span}(v_1, \dots, v_n)$ contains v_1 and each $v_k - v_1 = z_{k-1}$, so it equals V, making $v_1, \dots, v_n$ a basis of V (by 2.42). Finally

$$Tv_k = Tv_1 + Tz_{k-1} = w = w_1 + \cdots + w_m$$

for $2 \le k \le n$, and $Tv_1 = w$ as well; so all entries of $\mathcal{M}(T)$ with respect to these bases equal 1.

Problem 3C.3. Suppose $v_1, \dots, v_n$ is a basis of V and $w_1, \dots, w_m$ is a basis of W.
- (a) Show that if $S, T \in \mathcal{L}(V, W)$, then $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.
- (b) Show that if $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$, then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

[This exercise asks you to verify 3.35 and 3.38.]

Solution. Both identities follow by reading off coefficients in column k, which are unique because $w_1, \dots, w_m$ is a basis (2.28). Write $A = \mathcal{M}(S)$, $C = \mathcal{M}(T)$, and fix $k \in \{1, \dots, n\}$.
(a) Since $S + T \in \mathcal{L}(V, W)$ (3.5, 3.6),

$$\begin{aligned}(S+T)v_k = Sv_k + Tv_k &= \sum_{j=1}^{m} A_{j,k}\, w_j + \sum_{j=1}^{m} C_{j,k}\, w_j \\ &= \sum_{j=1}^{m} (A_{j,k} + C_{j,k})\, w_j,\end{aligned}$$

so $\mathcal{M}(S+T)_{j,k} = A_{j,k} + C_{j,k} = (A + C)_{j,k}$ by 3.34. Hence $\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T)$.
(b) Since $\lambda T \in \mathcal{L}(V, W)$ (3.5, 3.6),

$$(\lambda T)v_k = \lambda(Tv_k) = \lambda \sum_{j=1}^{m} C_{j,k}\, w_j = \sum_{j=1}^{m} (\lambda C_{j,k})\, w_j,$$

so $\mathcal{M}(\lambda T)_{j,k} = \lambda C_{j,k} = (\lambda C)_{j,k}$ by 3.36. Hence $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

Problem 3C.4. Suppose that $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ is the differentiation map defined by $Dp = p'$. Find a basis of $\mathcal{P}_3(\mathbf{R})$ and a basis of $\mathcal{P}_2(\mathbf{R})$ such that the matrix of D with respect to these bases is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

[Compare with Example 3.33. The next exercise generalizes this exercise.]

Solution. Take the basis of $\mathcal{P}_3(\mathbf{R})$ to be

$$v_1 = x, \qquad v_2 = \tfrac{x^2}{2}, \qquad v_3 = \tfrac{x^3}{3}, \qquad v_4 = 1,$$

and the basis of $\mathcal{P}_2(\mathbf{R})$ to be the standard one,

$$w_1 = 1, \qquad w_2 = x, \qquad w_3 = x^2.$$

The list v_1, v_2, v_3, v_4 spans $\mathcal{P}_3(\mathbf{R})$, since its span contains the scalar multiples $1, x, x^2, x^3$; having length $4 = \dim \mathcal{P}_3(\mathbf{R})$, it is a basis by 2.42. Differentiating,

$$\begin{aligned} Dv_1 &= (x)' = 1 && = 1w_1 + 0w_2 + 0w_3, \\ Dv_2 &= \left(\tfrac{x^2}{2}\right)' = x && = 0w_1 + 1w_2 + 0w_3, \\ Dv_3 &= \left(\tfrac{x^3}{3}\right)' = x^2 && = 0w_1 + 0w_2 + 1w_3, \\ Dv_4 &= (1)' = 0 && = 0w_1 + 0w_2 + 0w_3. \end{aligned}$$

Reading the coefficients off as columns,

$$\mathcal{M}(D) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Problem 3C.5. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ are 0 except that the entries in row k, column k, equal 1 if $1 \le k \le \dim \operatorname{range} T$.

Solution. Take the basis of V to be $v_1, \dots, v_r, u_1, \dots, u_{n-r}$, where $n = \dim V$, $r = \dim \operatorname{range} T$, the list $u_1, \dots, u_{n-r}$ is a basis of $\operatorname{null} T$ (its length is $n - r$ by 3.21), and $v_1, \dots, v_r$ are vectors extending it to a basis of V (2.32), reordered to put the v's first. The proof of 3.21 shows $Tv_1, \dots, Tv_r$ is a basis of $\operatorname{range} T$; set

$$w_k = Tv_k \quad (1 \le k \le r),$$

a linearly independent list in W, and extend it by 2.32 to a basis $w_1, \dots, w_m$ of W, where $m = \dim W$. With respect to these bases, column k of $A = \mathcal{M}(T)$ is read off from

$$Tv_k = w_k \quad (1 \le k \le r), \qquad Tu_i = 0 \quad (1 \le i \le n - r),$$

so $A_{k,k} = 1$ for $1 \le k \le r$ and every other entry of A is 0, as required.

Problem 3C.6. Suppose $v_1, \dots, v_m$ is a basis of V and W is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $w_1, \dots, w_n$ of W such that all entries in the first column of $\mathcal{M}(T)$ [with respect to the bases $v_1, \dots, v_m$ and $w_1, \dots, w_n$] are 0 except for possibly a 1 in the first row, first column.
[In this exercise, unlike Exercise 5, you are given the basis of V instead of being able to choose a basis of V.]

Solution. Choose the basis of W to have Tv_1 as its first vector when $Tv_1 \neq 0$. The first column of $A = \mathcal{M}(T)$ holds the coefficients in $Tv_1 = A_{1,1}w_1 + \cdots + A_{n,1}w_n$, unique by 2.28, so with $n = \dim W$:
(i) $Tv_1 = 0$. Let $w_1, \dots, w_n$ be any basis of W (one exists by 2.31). Then $Tv_1 = 0 \cdot w_1 + \cdots + 0 \cdot w_n$, so the first column is entirely 0 – permitted, since the 1 is only possibly there.
(ii) $Tv_1 \neq 0$. Then the one-vector list $w_1 = Tv_1$ is linearly independent, so it extends to a basis $w_1, \dots, w_n$ of W (2.32). Now

$$Tv_1 = w_1 = 1 \cdot w_1 + 0 \cdot w_2 + \cdots + 0 \cdot w_n,$$

so $A_{1,1} = 1$ and $A_{j,1} = 0$ for $j \ge 2$.

Problem 3C.7. Suppose $w_1, \ldots, w_n$ is a basis of W and V is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $v_1, \ldots, v_m$ of V such that all entries in the first row of $\mathcal{M}(T)$ [with respect to the bases $v_1, \ldots, v_m$ and $w_1, \ldots, w_n$] are 0 except for possibly a 1 in the first row, first column.
[In this exercise, unlike Exercise 5, you are given the basis of W instead of being able to choose a basis of W.]

Solution. Take v_1 with $\varphi(v_1) = 1$ and $v_2, \ldots, v_m$ a basis of $\operatorname{null} \varphi$, where $\varphi = \pi \circ T$ and $\pi \colon W \to \mathbf{F}$ sends $a_1 w_1 + \cdots + a_n w_n$ to a_1. Here π is linear because the coefficients of a vector in the basis $w_1, \ldots, w_n$ are unique and depend additively and homogeneously on the vector (2.28), so $\varphi \in \mathcal{L}(V, \mathbf{F})$ by 3.7. Writing $m = \dim V$ and $A = \mathcal{M}(T)$ with respect to $v_1, \ldots, v_m$ and $w_1, \ldots, w_n$, applying π to $T v_k = \sum_{j=1}^n A_{j,k} w_j$ gives

$$\varphi(v_k) = A_{1,k},$$

so the first row of $\mathcal{M}(T)$ is the list $\varphi(v_1), \ldots, \varphi(v_m)$.
(i) $\varphi = 0$. Any basis of V (2.31) makes the entire first row 0, which is permitted.
(ii) $\varphi \neq 0$. Then $\operatorname{range} \varphi$ is a nonzero subspace of the 1-dimensional space $\mathbf{F}$, hence equals $\mathbf{F}$ (2.37, 2.39), so $\dim \operatorname{null} \varphi = m - 1$ by 3.21 and $\varphi(v_1) = 1$ is attainable. The list $v_1, \ldots, v_m$ above is linearly independent: applying φ to $c_1 v_1 + \cdots + c_m v_m = 0$ gives $c_1 = 0$, and then independence of $v_2, \ldots, v_m$ gives $c_2 = \cdots = c_m = 0$. Having length $m = \dim V$, it is a basis of V (2.38), and

$$A_{1,1} = \varphi(v_1) = 1, \qquad A_{1,k} = \varphi(v_k) = 0 \ \text{ for } k \geq 2.$$

Problem 3C.8. Suppose A is an m-by-n matrix and B is an n-by-p matrix. Prove that

$$(AB)_{j,\cdot} = A_{j,\cdot} B$$

for each $1 \leq j \leq m$. In other words, show that row j of AB equals (row j of A) times B.
[This exercise gives the row version of 3.48.]

Solution. Both sides are 1-by-p matrices (by 3.41 and 3.44), so it suffices to compare entries in column k. Fix j and k; since $(A_{j,\cdot})_{1,r} = A_{j,r}$, two applications of 3.41 give

$$\left(A_{j,\cdot} B\right)_{1,k} = \sum_{r=1}^n \left(A_{j,\cdot}\right)_{1,r} B_{r,k} = \sum_{r=1}^n A_{j,r} B_{r,k} = (AB)_{j,k} = \left((AB)_{j,\cdot}\right)_{1,k}.$$

Hence $(AB)_{j,\cdot} = A_{j,\cdot} B$.

Problem 3C.9. Suppose $a = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ is a 1-by-n matrix and B is an n-by-p matrix. Prove that

$$aB = a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot}.$$

In other words, show that aB is a linear combination of the rows of B, with the scalars that multiply the rows coming from a.
[This exercise gives the row version of 3.50.]

Solution. Both sides are 1-by-p matrices (by 3.41, 3.44, 3.36, 3.34), so it suffices to compare entries in column k. Fix k; since the entry in row 1, column k, of $B_{r,\cdot}$ is $B_{r,k}$, the definitions of matrix multiplication (3.41), scalar multiplication (3.36), and matrix addition (3.34) give

$$(aB)_{1,k} = \sum_{r=1}^n a_r B_{r,k} = \left(a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot}\right)_{1,k}.$$

Hence $aB = a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot}$.

Problem 3C.10. Give an example of 2-by-2 matrices A and B such that $AB \neq BA$.

Solution. Take

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

By 3.41 (Check!),

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

which differ in row 1, column 1. Hence $AB \neq BA$.

Problem 3C.11. Prove that the distributive property holds for matrix addition and matrix multiplication. In other words, suppose A, B, C, D, E, and F are matrices whose sizes are such that $A(B+C)$ and $(D+E)F$ make sense. Explain why $AB+AC$ and $DF+EF$ both make sense and prove that

$$A(B+C) = AB + AC \qquad \text{and} \qquad (D+E)F = DF + EF.$$

Solution. The sizes force themselves: $A(B+C)$ making sense means B and C have a common size n-by-p (3.34) and A is m-by-n, so AB and AC are both defined and both m-by-p (3.41), hence addable, with $AB+AC$ of the same size m-by-p as $A(B+C)$. Likewise $(D+E)F$ making sense means D and E are both m-by-n and F is n-by-p, so DF, EF, and $DF+EF$ are all defined and m-by-p.
It therefore suffices to compare entries in row j, column k. Fix such j and k. By 3.41, then 3.34, then distributivity in $\mathbf{F}$ and splitting of a finite sum,

$$\begin{aligned}
\big(A(B+C)\big)_{j,k} &= \sum_{r=1}^{n} A_{j,r}\,(B+C)_{r,k} \\
&= \sum_{r=1}^{n} A_{j,r}\,\big(B_{r,k} + C_{r,k}\big) \\
&= \sum_{r=1}^{n} \big(A_{j,r}B_{r,k} + A_{j,r}C_{r,k}\big) \\
&= \sum_{r=1}^{n} A_{j,r}B_{r,k} + \sum_{r=1}^{n} A_{j,r}C_{r,k} \\
&= (AB)_{j,k} + (AC)_{j,k} \\
&= (AB+AC)_{j,k},
\end{aligned}$$

where the last two equalities use 3.41 and 3.34 again. Hence $A(B+C) = AB + AC$. The same three steps on the other side give

$$\begin{aligned}
\big((D+E)F\big)_{j,k} &= \sum_{r=1}^{n} \big(D_{j,r} + E_{j,r}\big)F_{r,k} \\
&= \sum_{r=1}^{n} D_{j,r}F_{r,k} + \sum_{r=1}^{n} E_{j,r}F_{r,k} = (DF+EF)_{j,k},
\end{aligned}$$

so $(D+E)F = DF + EF$.

Problem 3C.12. Prove that matrix multiplication is associative. In other words, suppose A, B, and C are matrices whose sizes are such that $(AB)C$ makes sense. Explain why $A(BC)$ makes sense and prove that

$$(AB)C = A(BC).$$

[Try to find a clean proof that illustrates the following quote from Emil Artin: “It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out.”]

Solution. Associativity of matrix multiplication is associativity of composition, transported through 3.43. First the sizes: $(AB)C$ making sense forces A to be m-by-n, B to be n-by-p, and C to be p-by-q, so by 3.41 the products BC (n-by-q) and then $A(BC)$ (m-by-q) are defined, of the same size as $(AB)C$. Throw the matrices out: for an m-by-n matrix A define $T_A \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ by

$$T_A(x_1, \dots, x_n) = \Big(\sum_{r=1}^{n} A_{1,r}x_r,\ \dots,\ \sum_{r=1}^{n} A_{m,r}x_r\Big),$$

which is linear because each coordinate is a fixed linear combination of $x_1, \dots, x_n$ (Check!). Sending the standard basis vector e_k to $T_A e_k = \sum_{j=1}^{m} A_{j,k} f_j$ exhibits column k of A as column k of $\mathcal{M}(T_A)$, so $\mathcal{M}(T_A) = A$ with respect to standard bases (3.31).

Form T_A, T_B, T_C this way. Two applications of 3.43, legitimate because standard bases are used throughout, give

$$\mathcal{M}\big((T_A T_B)T_C\big) = (AB)C, \qquad \mathcal{M}\big(T_A(T_B T_C)\big) = A(BC).$$

Composition of functions is associative, so $(T_A T_B)T_C = T_A(T_B T_C)$; equal maps have equal matrices, whence $(AB)C = A(BC)$.

Method (2): fix $j \in \{1, \dots, m\}$ and $k \in \{1, \dots, q\}$ and compute entries. By 3.41 twice, then interchanging two finite sums,

$$\begin{aligned}
\big((AB)C\big)_{j,k} &= \sum_{s=1}^{p} (AB)_{j,s} C_{s,k} = \sum_{s=1}^{p} \Big(\sum_{r=1}^{n} A_{j,r} B_{r,s}\Big) C_{s,k} = \sum_{s=1}^{p} \sum_{r=1}^{n} A_{j,r} B_{r,s} C_{s,k} \\
&= \sum_{r=1}^{n} \sum_{s=1}^{p} A_{j,r} B_{r,s} C_{s,k} = \sum_{r=1}^{n} A_{j,r} \Big(\sum_{s=1}^{p} B_{r,s} C_{s,k}\Big) = \sum_{r=1}^{n} A_{j,r} (BC)_{r,k} \\
&= \big(A(BC)\big)_{j,k}.
\end{aligned}$$

Since j and k were arbitrary, $(AB)C = A(BC)$.

Problem 3C.13. Suppose A is an n-by-n matrix and $1 \le j, k \le n$. Show that the entry in row j, column k, of A^3 (which is defined to mean AAA) is

$$\sum_{p=1}^{n} \sum_{r=1}^{n} A_{j,p} A_{p,r} A_{r,k}.$$

Solution. Group $A^3 = A^2 A$ (legitimate: matrix multiplication is associative by Exercise 3C.12, and every product here is n-by-n by 3.41) and apply 3.41 twice. Fix j and k; then

$$\begin{aligned}
(A^3)_{j,k} = (A^2 A)_{j,k} &= \sum_{r=1}^{n} (A^2)_{j,r}\, A_{r,k} = \sum_{r=1}^{n} \Big(\sum_{p=1}^{n} A_{j,p} A_{p,r}\Big) A_{r,k} \\
&= \sum_{r=1}^{n} \sum_{p=1}^{n} A_{j,p} A_{p,r} A_{r,k} = \sum_{p=1}^{n} \sum_{r=1}^{n} A_{j,p} A_{p,r} A_{r,k},
\end{aligned}$$

the third equality by distributivity in $\mathbf{F}$ and the last by interchanging two finite sums.

Problem 3C.14. Suppose m and n are positive integers. Prove that the function $A \mapsto A^{\mathrm{t}}$ is a linear map from $\mathbf{F}^{m,n}$ to $\mathbf{F}^{n,m}$.

Solution. Write $\Phi(A) = A^{\mathrm{t}}$, an n-by-m matrix by 3.54, and check the two conditions of 3.1 entrywise: both sides of each identity are n-by-m matrices, so it suffices to fix $j \in \{1, \ldots, m\}$, $k \in \{1, \ldots, n\}$ and compare entries in row k, column j.
Additivity. For $A, B \in \mathbf{F}^{m,n}$, using 3.54 for the outer equalities and 3.34 for the inner ones,

$$\big((A+B)^{\mathrm{t}}\big)_{k,j} = (A+B)_{j,k} = A_{j,k} + B_{j,k} = \big(A^{\mathrm{t}} + B^{\mathrm{t}}\big)_{k,j},$$

so $\Phi(A+B) = \Phi(A) + \Phi(B)$.
Homogeneity. For $\lambda \in \mathbf{F}$, the same two citations with 3.36 in place of 3.34 give

$$\big((\lambda A)^{\mathrm{t}}\big)_{k,j} = (\lambda A)_{j,k} = \lambda A_{j,k} = \big(\lambda A^{\mathrm{t}}\big)_{k,j},$$

so $\Phi(\lambda A) = \lambda\, \Phi(A)$. Hence $\Phi \in \mathcal{L}(\mathbf{F}^{m,n}, \mathbf{F}^{n,m})$.

Problem 3C.15. Prove that if A is an m-by-n matrix and C is an n-by-p matrix, then

$$(AC)^t = C^t A^t.$$

This exercise shows that the transpose of the product of two matrices is the product of the transposes in the opposite order.

Solution. Both sides are p-by-m matrices – AC is m-by-p by 3.41 and transposes by 3.54, while C^t is p-by-n and A^t is n-by-m, so $C^t A^t$ is defined – so it suffices to compare entries in row k, column j. Fix $k \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, m\}$; using 3.54 to strip transposes and 3.41 for each product,

$$\begin{aligned}\big((AC)^t\big)_{k,j} = (AC)_{j,k} &= \sum_{r=1}^{n} A_{j,r} C_{r,k} = \sum_{r=1}^{n} (C^t)_{k,r} (A^t)_{r,j} \\ &= (C^t A^t)_{k,j},\end{aligned}$$

the middle equality because multiplication in $\mathbf{F}$ is commutative. Hence $(AC)^t = C^t A^t$.

Problem 3C.16. Suppose A is an m-by-n matrix with $A \neq 0$. Prove that the rank of A is 1 if and only if there exist $(c_1, \ldots, c_m) \in \mathbf{F}^m$ and $(d_1, \ldots, d_n) \in \mathbf{F}^n$ such that

$$A_{j,k} = c_j d_k$$

for every $j = 1, \ldots, m$ and every $k = 1, \ldots, n$.

Solution. Both conditions say that the columns of A all lie on one line: write $U = \operatorname{span}(A_{\cdot,1}, \ldots, A_{\cdot,n}) \subseteq \mathbf{F}^{m,1}$, so that $\dim U$ is the rank of A (by 3.52 and 3.58).
(i) Rank 1 implies the factorization. Then $\dim U = 1$, so $U = \operatorname{span}(c)$ for some nonzero $c \in \mathbf{F}^{m,1}$, say with entries $c_1, \ldots, c_m$. Each column $A_{\cdot,k}$ lies in U, so $A_{\cdot,k} = d_k c$ for some $d_k \in \mathbf{F}$; comparing row j (using 3.44) gives

$$A_{j,k} = d_k c_j = c_j d_k \qquad (1 \le j \le m,\ 1 \le k \le n).$$

(ii) The factorization implies rank 1. Let $c \in \mathbf{F}^{m,1}$ have entries $c_1, \ldots, c_m$. Since row j of $A_{\cdot,k}$ is $A_{j,k} = c_j d_k$, which is row j of $d_k c$, every column satisfies $A_{\cdot,k} = d_k c$, so

$$U \subseteq \operatorname{span}(c), \qquad \dim U \le \dim \operatorname{span}(c) \le 1$$

by 2.37 (a one-element list spans a space of dimension at most 1, by 2.30). Since $A \neq 0$, some column $A_{\cdot,k}$ is a nonzero element of U, so $\dim U \ge 1$. Hence $\dim U = 1$.

Problem 3C.17. Suppose $T \in \mathcal{L}(V)$, and $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are bases of V. Prove that the following are equivalent.
(a) T is injective.
(b) The columns of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{n,1}$.
(c) The columns of $\mathcal{M}(T)$ span $\mathbf{F}^{n,1}$.
(d) The rows of $\mathcal{M}(T)$ span $\mathbf{F}^{1,n}$.
(e) The rows of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{1,n}$.
Here $\mathcal{M}(T)$ means $\mathcal{M}(T, (u_1, \ldots, u_n), (v_1, \ldots, v_n))$.

Solution. Everything rests on the dictionary $x \mapsto u = x_1u_1 + \cdots + x_nu_n$ from $\mathbf{F}^{n,1}$ onto V, under which $Tu = 0$ exactly when $Ax = 0$. Here $A = \mathcal{M}(T)$ is n-by-n (both lists are bases, so $\dim V = n$) and $Tu_k = \sum_j A_{j,k}v_j$ by 3.31, whence linearity of T gives

$$Tu = \sum_{k=1}^{n} x_k\, Tu_k = \sum_{j=1}^{n} (Ax)_{j,1}\, v_j, \qquad Ax = x_1A_{\cdot,1} + \cdots + x_nA_{\cdot,n},$$

the second identity by 3.50. Since $v_1, \ldots, v_n$ is linearly independent, $Tu = 0$ iff $Ax = 0$; since $u_1, \ldots, u_n$ is a basis, $u = 0$ iff $x = 0$, and every $u \in V$ arises from some x.
(a) $\iff$ (b). By 3.15, T is injective iff $\operatorname{null} T = \{0\}$, which by the dictionary holds iff the only x with $x_1A_{\cdot,1} + \cdots + x_nA_{\cdot,n} = Ax = 0$ is $x = 0$ – that is, iff the columns of A are linearly independent.
(b) $\iff$ (c). The columns form a list of length n in $\mathbf{F}^{n,1}$, and $\dim \mathbf{F}^{n,1} = n$ (3.40); such a list is a basis as soon as it is linearly independent (2.38) or spanning (2.42), so each of (b), (c) implies the other.
(c) $\iff$ (d). Writing U for the span of the columns, (c) says $U = \mathbf{F}^{n,1}$, which by 2.39 and 3.40 holds iff $\dim U = n$, i.e. iff the column rank of A is n (3.52, 3.58). The same argument in $\mathbf{F}^{1,n}$ shows (d) holds iff the row rank of A is n. The two ranks are equal by 3.57.
(d) $\iff$ (e). As in (b) $\iff$ (c), applied to the n rows in $\mathbf{F}^{1,n}$.

3.4 Exercises 3D

Problem 3D.1. Suppose $T \in \mathcal{L}(V, W)$ is invertible. Show that T^{-1} is invertible and

$$(T^{-1})^{-1} = T.$$

Solution. Take $T \in \mathcal{L}(V, W)$ itself as the inverse of $T^{-1} \in \mathcal{L}(W, V)$. Invertibility of T (3.59) gives

$$TT^{-1} = I_W \qquad \text{and} \qquad T^{-1}T = I_V,$$

which are precisely the two identities 3.59 demands of an inverse of T^{-1}. Hence T^{-1} is invertible, and $(T^{-1})^{-1} = T$ by uniqueness of the inverse (3.60).

Problem 3D.2. Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$ are both invertible linear maps. Prove that $ST \in \mathcal{L}(U, W)$ is invertible and that $(ST)^{-1} = T^{-1}S^{-1}$.

Solution. Take $T^{-1}S^{-1} \in \mathcal{L}(W, U)$ as the inverse; it and $ST \in \mathcal{L}(U, W)$ are linear because compositions of linear maps are linear (3.8). Associativity of composition gives

$$(T^{-1}S^{-1})(ST) = T^{-1}(S^{-1}S)T = T^{-1}I_VT = T^{-1}T = I_U,$$

$$(ST)(T^{-1}S^{-1}) = S(TT^{-1})S^{-1} = SI_VS^{-1} = SS^{-1} = I_W.$$

Hence ST is invertible (3.59), and $(ST)^{-1} = T^{-1}S^{-1}$ by uniqueness of the inverse (3.60).

Problem 3D.3. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent.
(a) T is invertible.
(b) $Tv_1, \ldots, Tv_n$ is a basis of V for every basis $v_1, \ldots, v_n$ of V.
(c) $Tv_1, \ldots, Tv_n$ is a basis of V for some basis $v_1, \ldots, v_n$ of V.

Solution. We run the cycle (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a), writing $n = \dim V$.
(a) $\Rightarrow$ (b). Let $v_1, \ldots, v_n$ be a basis of V. If $a_1Tv_1 + \cdots + a_nTv_n = 0$, then $T(a_1v_1 + \cdots + a_nv_n) = 0$, and T is injective (3.63) with $\operatorname{null} T = \{0\}$ (3.15), so $a_1v_1 + \cdots + a_nv_n = 0$ and all $a_k = 0$; thus $Tv_1, \ldots, Tv_n$ is linearly independent. It spans V because for $v \in V$, writing $T^{-1}v = a_1v_1 + \cdots + a_nv_n$ and applying T gives

$$v = T(T^{-1}v) = a_1Tv_1 + \cdots + a_nTv_n.$$

Hence $Tv_1, \ldots, Tv_n$ is a basis of V.
(b) $\Rightarrow$ (c). Apply (b) to a basis of V, which exists by 2.31.
(c) $\Rightarrow$ (a). By Exercise 10 in Section 3B, $\operatorname{range} T = \operatorname{span}(Tv_1, \ldots, Tv_n) = V$, the last equality because $Tv_1, \ldots, Tv_n$ is a basis of V. So T is surjective, hence injective (3.65, as V is finite-dimensional), hence invertible (3.63).

Problem 3D.4. Suppose V is finite-dimensional and $\dim V > 1$. Prove that the set of noninvertible linear maps from V to itself is not a subspace of $\mathcal{L}(V)$.

Solution. The set $\mathcal{N}$ of noninvertible operators is not closed under addition, hence is not a subspace (1.34). Fix a basis $v_1, \ldots, v_n$ of V (2.31), where $n = \dim V \geq 2$, and let $S, T \in \mathcal{L}(V)$ be given by the linear map lemma (3.4) via

$$Sv_1 = v_1, \quad Sv_k = 0 \ (k \geq 2); \qquad Tv_1 = 0, \quad Tv_k = v_k \ (k \geq 2).$$

Since $n \geq 2$, the basis vectors v_1, v_2 exist and are nonzero, while $Sv_2 = 0$ and $Tv_1 = 0$; so neither S nor T is injective (3.15), hence neither is invertible (3.63), and $S, T \in \mathcal{N}$. But $S + T$ agrees with I on the basis,

$$(S+T)v_1 = v_1 + 0 = v_1, \qquad (S+T)v_k = 0 + v_k = v_k \ \ (k \geq 2),$$

so $S + T = I$, which is invertible. Hence $S + T \notin \mathcal{N}$.

Problem 3D.5. Suppose V is finite-dimensional, U is a subspace of V, and $S \in \mathcal{L}(U, V)$. Prove that there exists an invertible linear map T from V to itself such that $Tu = Su$ for every $u \in U$ if and only if S is injective.

Solution. (i) Necessity. If such an invertible T exists and $u \in U$ has $Su = 0$, then $Tu = Su = 0$, and T is injective (3.63) with $\operatorname{null} T = \{0\}$ (3.15), so $u = 0$. Thus $\operatorname{null} S = \{0\}$ and S is injective (3.15).
(ii) Sufficiency. Suppose S is injective. Let $u_1, \ldots, u_m$ be a basis of the (finite-dimensional, by 2.25) subspace U and extend it to a basis $u_1, \ldots, u_m, v_1, \ldots, v_p$ of V (2.32). The list $Su_1, \ldots, Su_m$ is linearly independent: $a_1Su_1 + \cdots + a_mSu_m = 0$ gives $a_1u_1 + \cdots + a_mu_m \in \operatorname{null} S = \{0\}$, whence all $a_k = 0$. Extend it to a basis

$$Su_1, \ldots, Su_m, w_1, \ldots, w_q$$

of V (2.32); since every basis of V has length $\dim V$ (2.34), $q = p$. So the linear map lemma (3.4) supplies $T \in \mathcal{L}(V)$ with

$$Tu_k = Su_k \ \ (k = 1, \ldots, m), \qquad Tv_j = w_j \ \ (j = 1, \ldots, p).$$

This T carries a basis of V to a basis of V, so it is invertible by Exercise 3D.3 ((c) $\Rightarrow$ (a)); and for $u = a_1u_1 + \cdots + a_mu_m \in U$, linearity gives

$$Tu = a_1Su_1 + \cdots + a_mSu_m = S(a_1u_1 + \cdots + a_mu_m) = Su.$$

Problem 3D.6. Suppose that W is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} S = \operatorname{null} T$ if and only if there exists an invertible $E \in \mathcal{L}(W)$ such that $S = ET$.

Solution. (i) Necessity. If $S = ET$ with E invertible, then $Tv = 0$ gives $Sv = E0 = 0$, while $Sv = 0$ gives $E(Tv) = 0$ and hence $Tv = 0$, since E is injective (3.63) with $\operatorname{null} E = \{0\}$ (3.15). So $\operatorname{null} S = \operatorname{null} T$.
(ii) Sufficiency. Suppose $\operatorname{null} S = \operatorname{null} T$, and define $\varphi\colon \operatorname{range} T \to \operatorname{range} S$ by $\varphi(Tv) = Sv$. This is unambiguous precisely because of the hypothesis: $Tv_1 = Tv_2$ forces $v_1 - v_2 \in \operatorname{null} T = \operatorname{null} S$, hence $Sv_1 = Sv_2$. It is linear, since $\varphi(\lambda Tv_1 + Tv_2) = S(\lambda v_1 + v_2) = \lambda\varphi(Tv_1) + \varphi(Tv_2)$; surjective onto $\operatorname{range} S$ by construction; and injective, because $\varphi(Tv) = Sv = 0$ puts $v \in \operatorname{null} S = \operatorname{null} T$, so $Tv = 0$ (3.15). Thus φ is an isomorphism (3.63, 3.69).
Let $u_1, \dots, u_m$ be a basis of $\operatorname{range} T$ (finite-dimensional as a subspace of W, by 2.25) and extend it to a basis $u_1, \dots, u_n$ of W (2.32), where $n = \dim W$. The list $\varphi u_1, \dots, \varphi u_m$ is again linearly independent (apply the injective φ to a vanishing combination), so extend it to a basis

$$\varphi u_1, \dots, \varphi u_m, y_{m+1}, \dots, y_n$$

of W (2.32; the count $n - m$ because every basis of W has length n, by 2.34). The linear map lemma (3.4) now gives $E \in \mathcal{L}(W)$ with

$$Eu_k = \varphi u_k \quad (k \le m), \qquad Eu_k = y_k \quad (k > m),$$

which carries a basis of W to a basis of W and so is invertible by Exercise 3D.3 ((c) $\Rightarrow$ (a)). Finally, for $v \in V$ write $Tv = c_1 u_1 + \cdots + c_m u_m$; then linearity of E and of φ give

$$E(Tv) = c_1 \varphi u_1 + \cdots + c_m \varphi u_m = \varphi(Tv) = Sv,$$

so $S = ET$.

Problem 3D.7. Suppose that V is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} S = \operatorname{range} T$ if and only if there exists an invertible $E \in \mathcal{L}(V)$ such that $S = TE$.

Solution. (i) Necessity. If $S = TE$ with E invertible, then E is surjective (3.63), so $\operatorname{range} E = V$ and

$$\operatorname{range} S = \{T(Ev) : v \in V\} = \{Tu : u \in V\} = \operatorname{range} T.$$

(ii) Sufficiency. Suppose $\operatorname{range} S = \operatorname{range} T$; we build E from adapted bases. The fundamental theorem of linear maps (3.21), applicable since V is finite-dimensional, gives

$$\dim \operatorname{null} S = \dim V - \dim \operatorname{range} S = \dim \operatorname{null} T;$$

call this common value m and set $r = \dim V - m$. Take a basis $u_1, \dots, u_m$ of $\operatorname{null} S$ extended to a basis $u_1, \dots, u_m, v_1, \dots, v_r$ of V (2.32, 2.34), and a basis $p_1, \dots, p_m$ of $\operatorname{null} T$ (2.25, 2.31).
Then $Sv_1, \dots, Sv_r$ is a basis of $\operatorname{range} S$: it spans, since $Su_k = 0$ makes $Sv = a_1 Sv_1 + \cdots + a_r Sv_r$ for $v = \sum_k b_k u_k + \sum_j a_j v_j$; and if $\sum_j a_j Sv_j = 0$ then $\sum_j a_j v_j \in \operatorname{null} S = \operatorname{span}(u_1, \dots, u_m)$, so

$$b_1 u_1 + \cdots + b_m u_m - a_1 v_1 - \cdots - a_r v_r = 0$$

for some b_k, forcing every $a_j = 0$ by independence of the basis of V.
Since $Sv_j \in \operatorname{range} S = \operatorname{range} T$, choose $q_j \in V$ with $Tq_j = Sv_j$. The list $p_1, \dots, p_m, q_1, \dots, q_r$ has length $m + r = \dim V$, so by 2.38 it is a basis of V once independent: applying T to a vanishing combination kills the p_k and leaves $\sum_j a_j Sv_j = 0$, so all $a_j = 0$ by the previous paragraph, and then all $b_k = 0$. The linear map lemma (3.4) therefore gives $E \in \mathcal{L}(V)$ with

$$Eu_k = p_k \quad (k = 1, \dots, m), \qquad Ev_j = q_j \quad (j = 1, \dots, r),$$

invertible by Exercise 3D.3 ((c) $\Rightarrow$ (a)) since it carries a basis of V to a basis of V. Finally TE and S agree on the basis $u_1, \dots, u_m, v_1, \dots, v_r$, namely

$$(TE)u_k = Tp_k = 0 = Su_k, \qquad (TE)v_j = Tq_j = Sv_j,$$

so $S = TE$.

Problem 3D.8. Suppose V and W are finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that there exist invertible $E_1 \in \mathcal{L}(V)$ and $E_2 \in \mathcal{L}(W)$ such that $S = E_2 T E_1$ if and only if $\dim \operatorname{null} S = \dim \operatorname{null} T$.

Solution. (i) Necessity. If $S = E_2 T E_1$ with both E_i invertible, then for $v \in V$, using injectivity of E_2 (3.63) for the middle step,

$$Sv = 0 \iff TE_1 v = 0 \iff E_1 v \in \operatorname{null} T,$$

so $\operatorname{null} S = \{E_1^{-1} u : u \in \operatorname{null} T\}$. Thus E_1^{-1} restricted to $\operatorname{null} T$ is linear, injective and onto $\operatorname{null} S$, hence an isomorphism (3.63); both null spaces are finite-dimensional (2.25), so $\dim \operatorname{null} S = \dim \operatorname{null} T$ by 3.70.
(ii) Sufficiency. Let n be the common null-space dimension and $k = \dim V - n$, so that

$$\dim \operatorname{range} S = k = \dim \operatorname{range} T$$

by the fundamental theorem of linear maps (3.21). Extend a basis $u_1, \ldots, u_n$ of $\operatorname{null} S$ to a basis $u_1, \ldots, u_n, x_1, \ldots, x_k$ of V, and a basis $p_1, \ldots, p_n$ of $\operatorname{null} T$ to a basis $p_1, \ldots, p_n, y_1, \ldots, y_k$ of V (2.32). Since $Su_j = 0$, the list $Sx_1, \ldots, Sx_k$ spans $\operatorname{range} S$ and has length $k = \dim \operatorname{range} S$, so it is a basis of $\operatorname{range} S$ (2.42); likewise $Ty_1, \ldots, Ty_k$ is a basis of $\operatorname{range} T$. Extend both to bases of W (2.32),

$$Sx_1, \ldots, Sx_k, w_1, \ldots, w_r \qquad \text{and} \qquad Ty_1, \ldots, Ty_k, z_1, \ldots, z_r,$$

with the same $r = \dim W - k$ in each. The linear map lemma (3.4) now gives $E_1 \in \mathcal{L}(V)$ and $E_2 \in \mathcal{L}(W)$ with

$$E_1 u_j = p_j, \quad E_1 x_i = y_i; \qquad E_2(Ty_i) = Sx_i, \quad E_2 z_j = w_j.$$

Each carries a basis to a basis, hence is invertible by Exercise 3D.3 ((c) $\Rightarrow$ (a)). Finally S and $E_2 T E_1$ agree on the basis $u_1, \ldots, u_n, x_1, \ldots, x_k$ of V:

$$E_2 T E_1 u_j = E_2 T p_j = 0 = S u_j, \qquad E_2 T E_1 x_i = E_2 T y_i = S x_i,$$

so $S = E_2 T E_1$.

Problem 3D.9. Suppose V is finite-dimensional and $T \colon V \to W$ is a surjective linear map of V onto W. Prove that there is a subspace U of V such that $T|_U$ is an isomorphism of U onto W. Here $T|_U$ means the function T restricted to U. Thus $T|_U$ is the function whose domain is U, with $T|_U$ defined by $T|_U(u) = Tu$ for every $u \in U$.

Solution. Take any subspace U of V with $V = \operatorname{null} T \oplus U$, which exists by 2.33 since V is finite-dimensional and $\operatorname{null} T$ is a subspace of V. The restriction $T|_U \in \mathcal{L}(U, W)$ inherits additivity and homogeneity from T, and:
(i) Injective. If $u \in U$ and $Tu = 0$, then $u \in \operatorname{null} T \cap U = \{0\}$, the intersection being trivial because the sum is direct (1.46); so $\operatorname{null}(T|_U) = \{0\}$ and 3.15 applies.
(ii) Surjective. Given $w \in W$, surjectivity of T supplies $v \in V$ with $Tv = w$; writing $v = x + u$ with $x \in \operatorname{null} T$ and $u \in U$,

$$w = Tv = Tx + Tu = 0 + Tu = T|_U(u).$$

Hence $T|_U$ is invertible (3.63), i.e. an isomorphism of U onto W.

Problem 3D.10. Suppose V and W are finite-dimensional and U is a subspace of V. Let

$$\mathcal{E} = \{T \in \mathcal{L}(V, W) : U \subseteq \operatorname{null} T\}.$$

(a) Show that $\mathcal{E}$ is a subspace of $\mathcal{L}(V, W)$.

(b) Find a formula for $\dim \mathcal{E}$ in terms of $\dim V$, $\dim W$, and $\dim U$.
Hint: Define $\Phi\colon \mathcal{L}(V,W) \to \mathcal{L}(U,W)$ by $\Phi(T) = T|_U$. What is $\operatorname{null} \Phi$? What is $\operatorname{range} \Phi$?

Solution. (a) The zero map lies in $\mathcal{E}$ (its null space is V), and for $S, T \in \mathcal{E}$, $\lambda \in \mathbf{F}$, $u \in U$,

$$(S+T)u = 0+0 = 0, \qquad (\lambda T)u = \lambda 0 = 0,$$

so $S+T$ and λT lie in $\mathcal{E}$. Hence $\mathcal{E}$ is a subspace of $\mathcal{L}(V,W)$ by 1.34.
(b) The formula is

$$\dim \mathcal{E} = (\dim V - \dim U)(\dim W).$$

Following the hint, let $\Phi(T) = T|_U$ map $\mathcal{L}(V,W)$ to $\mathcal{L}(U,W)$; it is linear because restriction preserves the pointwise operations, $(S+T)|_U = S|_U + T|_U$ and $(\lambda T)|_U = \lambda(T|_U)$.
Its null space is $\mathcal{E}$, since $T|_U = 0$ says exactly $U \subseteq \operatorname{null} T$. It is surjective: given $R \in \mathcal{L}(U,W)$, extend a basis $u_1, \ldots, u_m$ of U (finite-dimensional by 2.25) to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V (2.32) and take the $T \in \mathcal{L}(V,W)$ given by the linear map lemma (3.4) with

$$Tu_j = Ru_j \quad (j = 1, \ldots, m), \qquad Tv_k = 0 \quad (k = 1, \ldots, n);$$

then $T|_U$ and R agree on a basis of U, so $\Phi(T) = R$. Since $\mathcal{L}(V,W)$ is finite-dimensional (3.72), the fundamental theorem of linear maps (3.21) applied to Φ gives

$$\dim \mathcal{E} = \dim \mathcal{L}(V,W) - \dim \mathcal{L}(U,W) = (\dim V - \dim U)(\dim W),$$

the last step evaluating both terms by 3.72.

Problem 3D.11. Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that

$$ST \text{ is invertible} \iff S \text{ and } T \text{ are invertible.}$$

Solution. ($\Longleftarrow$) If S and T are invertible then $T^{-1}S^{-1}$ is a two-sided inverse of ST (Check!), so ST is invertible.
($\Longrightarrow$) Suppose ST is invertible; then ST is injective and surjective (by 3.63).
(i) S is surjective: every $(ST)v = S(Tv)$ lies in $\operatorname{range} S$, so

$$V = \operatorname{range}(ST) \subseteq \operatorname{range} S \subseteq V.$$

Hence S is invertible, since a surjective operator on a finite-dimensional V is invertible (by 3.65).
(ii) T is injective: if $Tv = 0$ then $(ST)v = S0 = 0$, so $v = 0$ by injectivity of ST and 3.15. Hence T is invertible (by 3.65 again, V finite-dimensional).

Problem 3D.12. Suppose V is finite-dimensional and $S, T, U \in \mathcal{L}(V)$ and $STU = I$. Show that T is invertible and that $T^{-1} = US$.

Solution. Both S and U are invertible, and then $T = S^{-1}U^{-1}$ forces $T^{-1} = US$.
(i) U is invertible: if $Uv = 0$ then $v = (STU)v = S(T0) = 0$, so U is injective (by 3.15) and hence invertible, V being finite-dimensional (by 3.65).
(ii) S is invertible: every $(STU)v = S\big((TU)v\big)$ lies in $\operatorname{range} S$, so

$$V = \operatorname{range} I = \operatorname{range}(STU) \subseteq \operatorname{range} S \subseteq V,$$

making S surjective and hence invertible (by 3.65).
Multiplying $STU = I$ by S^{-1} on the left and U^{-1} on the right gives $T = S^{-1}U^{-1}$, a product of invertible maps, hence invertible with

$$T^{-1} = (S^{-1}U^{-1})^{-1} = (U^{-1})^{-1}(S^{-1})^{-1} = US,$$

by Exercises 1 and 2 in this section.

Problem 3D.13. Show that the result in Exercise 12 can fail without the hypothesis that V is finite-dimensional.

Solution. On $V = \mathbf{F}^\infty$ take $S = T = B$ and $U = F^2$, where B and F are the backward and forward shifts

$$B(x_1, x_2, x_3, \dots) = (x_2, x_3, x_4, \dots), \quad F(x_1, x_2, \dots) = (0, x_1, x_2, \dots),$$

both linear by Example 3.3. Then $STU = B^2F^2 = I$, since

$$B^2F^2(x_1, x_2, \dots) = B^2(0, 0, x_1, x_2, \dots) = (x_1, x_2, \dots).$$

But $T = B$ is not invertible: $B(1, 0, 0, \dots) = 0$, so B is not injective (by 3.15) and hence not invertible (by 3.63).

Problem 3D.14. Prove or give a counterexample: If V is a finite-dimensional vector space and $R, S, T \in \mathcal{L}(V)$ are such that RST is surjective, then S is injective.

Solution. True: S is even invertible. Because V is finite-dimensional and the operator RST is surjective, RST is invertible (by 3.65). Writing $RST = R(ST)$, Exercise 11 gives that R and ST are invertible; applying Exercise 11 to the invertible ST gives that S (and T) is invertible. In particular S is injective (by 3.63).

Problem 3D.15. Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_m$ is a list in V such that $Tv_1, \dots, Tv_m$ spans V. Prove that $v_1, \dots, v_m$ spans V.

Solution. T is injective, and injectivity converts the spanning of $Tv_1, \dots, Tv_m$ into that of $v_1, \dots, v_m$. Indeed V is finite-dimensional, being spanned by the finite list $Tv_1, \dots, Tv_m$, and each Tv_k lies in the subspace $\operatorname{range} T$ (by 3.18), so

$$V = \operatorname{span}(Tv_1, \dots, Tv_m) \subseteq \operatorname{range} T \subseteq V;$$

thus T is surjective and hence injective (by 3.65, V finite-dimensional).
Now let $v \in V$. Because $Tv_1, \dots, Tv_m$ spans V, there are $a_1, \dots, a_m \in \mathbf{F}$ with

$$Tv = a_1Tv_1 + \cdots + a_mTv_m = T(a_1v_1 + \cdots + a_mv_m),$$

so injectivity gives $v = a_1v_1 + \cdots + a_mv_m$. Hence $v_1, \dots, v_m$ spans V.

Problem 3D.16. Prove that every linear map from $\mathbf{F}^{n,1}$ to $\mathbf{F}^{m,1}$ is given by a matrix multiplication. In other words, prove that if $T \in \mathcal{L}(\mathbf{F}^{n,1}, \mathbf{F}^{m,1})$, then there exists an m-by-n matrix A such that $Tx = Ax$ for every $x \in \mathbf{F}^{n,1}$.

Solution. Take A to be the m-by-n matrix whose k^{th} column is Te_k, where $e_1, \dots, e_n$ is the standard basis of $\mathbf{F}^{n,1}$ (so e_k has k^{th} entry 1 and other entries 0); that is, $A_{j,k} = (Te_k)_j$. For $x \in \mathbf{F}^{n,1}$ with entries $x_1, \dots, x_n$ we have $x = x_1e_1 + \cdots + x_ne_n$, so linearity of T gives

$$\begin{aligned} Tx &= x_1Te_1 + \cdots + x_nTe_n \\ &= x_1A_{\cdot,1} + \cdots + x_nA_{\cdot,n} = Ax, \end{aligned}$$

the last equality being 3.50 (Ax is that linear combination of the columns of A).

Problem 3D.17. Suppose V is finite-dimensional and $S \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(T) = ST$$

for $T \in \mathcal{L}(V)$.
(a) Show that $\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S)$.
(b) Show that $\dim \operatorname{range} \mathcal{A} = (\dim V)(\dim \operatorname{range} S)$.

Solution. Write $n = \dim V$ and $U = \operatorname{null} S$, a finite-dimensional subspace of V (by 2.25); by 3.72, $\dim \mathcal{L}(V) = n^2$.
(a) The null space of $\mathcal{A}$ is the set of T with range inside U:

$$\operatorname{null} \mathcal{A} = \{T \in \mathcal{L}(V) : \operatorname{range} T \subseteq U\},$$

since $ST = 0$ says $Tv \in \operatorname{null} S$ for every $v \in V$. Enlarging the target space from U to V gives a map $\Phi : \mathcal{L}(V, U) \to \mathcal{L}(V)$, $\big(\Phi(R)\big)v = Rv$, which is linear and injective (if $\Phi(R) = 0$ then $Rv = 0$ for all v), with range exactly $\operatorname{null} \mathcal{A}$: each $\Phi(R)$ has range in U, and conversely any such T is $\Phi(R)$ for $R \in \mathcal{L}(V, U)$ defined by $Rv = Tv$. So Φ is an isomorphism onto $\operatorname{null} \mathcal{A}$, and 3.72 (both V and U finite-dimensional) gives

$$\dim \operatorname{null} \mathcal{A} = \dim \mathcal{L}(V, U) = (\dim V)(\dim \operatorname{null} S).$$

(b) Apply 3.21 to $\mathcal{A}$ on the finite-dimensional space $\mathcal{L}(V)$, then to S on V:

$$\begin{aligned}\dim \operatorname{range} \mathcal{A} &= n^2 - n\,(\dim \operatorname{null} S) = n\big(n - \dim \operatorname{null} S\big)\\ &= (\dim V)(\dim \operatorname{range} S).\end{aligned}$$

Problem 3D.18. Show that V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces.

Solution. Evaluation at 1 is an isomorphism: define $\Phi : \mathcal{L}(\mathbf{F}, V) \to V$ by $\Phi(T) = T1$. Linearity of Φ is immediate from the definitions of addition and scalar multiplication in $\mathcal{L}(\mathbf{F}, V)$ (Check!).
Φ is injective: if $T1 = 0$ then homogeneity of T gives

$$T\lambda = T(\lambda \cdot 1) = \lambda T1 = 0 \qquad \text{for all } \lambda \in \mathbf{F},$$

so $T = 0$ and $\operatorname{null} \Phi = \{0\}$ (by 3.15).
Φ is surjective: given $v \in V$, the map $T_v\lambda = \lambda v$ is linear (Check!) and $\Phi(T_v) = 1 \cdot v = v$.
Hence Φ is invertible (by 3.63), so $\mathcal{L}(\mathbf{F}, V)$ and V are isomorphic (3.69).

Problem 3D.19. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has the same matrix with respect to every basis of V if and only if T is a scalar multiple of the identity operator.

Solution. ($\Longleftarrow$) If $T = \lambda I$ then $Tv_k = \lambda v_k$ for every basis $v_1, \ldots, v_n$, so $\mathcal{M}(T)$ is the diagonal matrix λI no matter which basis is used.
($\Longrightarrow$) Suppose $\mathcal{M}(T, (w_1, \ldots, w_n)) = A$ for every basis $w_1, \ldots, w_n$, and let $n = \dim V \geq 1$ (if $n = 0$ then $T = 0 = 0I$). Fix a basis $v_1, \ldots, v_n$.
(i) A is diagonal. Fix k and set $u_k = 2v_k$ and $u_i = v_i$ for $i \neq k$; this is again a basis (Check!; here $2 \neq 0$ in $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$). Column k of A read in each of the two bases gives

$$Tv_k = \sum_j A_{j,k} v_j \quad \text{and} \quad 2\,Tv_k = Tu_k = 2A_{k,k}v_k + \sum_{j \neq k} A_{j,k} v_j.$$

Twice the first minus the second yields $\sum_{j \neq k} A_{j,k} v_j = 0$, so $A_{j,k} = 0$ for $j \neq k$ by linear independence. Write $\lambda_k = A_{k,k}$; then $Tw_k = \lambda_k w_k$ for every basis $w_1, \ldots, w_n$ and every k.

(ii) The λ_k agree. For $k \geq 2$, swapping v_1 and v_k gives a basis with first vector v_k, so $Tv_k = \lambda_1 v_k$; comparing with $Tv_k = \lambda_k v_k$ and $v_k \neq 0$ gives $\lambda_k = \lambda_1 =: \lambda$.
Thus T and λI agree on the basis $v_1, \dots, v_n$, hence $T = \lambda I$ (by 3.4).

Problem 3D.20. Suppose $q \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{R})$ such that

$$q(x) = (x^2 + x)p''(x) + 2x\,p'(x) + p(3)$$

for all $x \in \mathbf{R}$.

Solution. Fix m with $q \in \mathcal{P}_m(\mathbf{R})$ (take $m = \deg q$, or $m = 0$ if $q = 0$) and define $T : \mathcal{P}_m(\mathbf{R}) \to \mathcal{P}_m(\mathbf{R})$ by

$$(Tp)(x) = (x^2 + x)p''(x) + 2x\,p'(x) + p(3).$$

This does land in $\mathcal{P}_m(\mathbf{R})$, since $\deg p \leq m$ gives $\deg\big((x^2 + x)p''\big) \leq m$ and $\deg(2xp') \leq m$ while $p(3)$ is constant; and T is linear because differentiation, evaluation at 3, and multiplication by a fixed polynomial are all linear (Check!).
T is injective: let $p \neq 0$ have degree d and leading coefficient $a_d \neq 0$.
(i) $d = 0$: then $p' = p'' = 0$, so $Tp = p(3) = a_0 \neq 0$.
(ii) $d \geq 1$: as $p(3)$ contributes only to the constant term, the coefficient of x^d in Tp is

$$d(d-1)a_d + 2d\,a_d = a_d\,d(d+1) \neq 0,$$

so $Tp \neq 0$. (For $d = 1$ the first term is 0, matching $p'' = 0$; for $d \geq 2$ it is the $x^2 \cdot d(d-1)a_d x^{d-2}$ term.)
Hence $\operatorname{null} T = \{0\}$, so T is injective (by 3.15) and therefore surjective, being an operator on the finite-dimensional $\mathcal{P}_m(\mathbf{R})$ (by 3.65). Thus some $p \in \mathcal{P}_m(\mathbf{R}) \subseteq \mathcal{P}(\mathbf{R})$ satisfies $Tp = q$, which is the asserted identity.

Problem 3D.21. Suppose n is a positive integer and $A_{j,k} \in \mathbf{F}$ for all $j, k = 1, \dots, n$. Prove that the following are equivalent (note that in both parts below, the number of equations equals the number of variables).
(a) The trivial solution $x_1 = \cdots = x_n = 0$ is the only solution to the homogeneous system of equations

$$\sum_{k=1}^{n} A_{1,k}x_k = 0$$

$$\vdots$$

$$\sum_{k=1}^{n} A_{n,k}x_k = 0.$$

(b) For every $c_1, \dots, c_n \in \mathbf{F}$, there exists a solution to the system of equations

$$\sum_{k=1}^{n} A_{1,k}x_k = c_1$$

$$\vdots$$

$$\sum_{k=1}^{n} A_{n,k}x_k = c_n.$$

Solution. Both statements are about the operator $T \in \mathcal{L}(\mathbf{F}^n)$ defined by

$$T(x_1, \dots, x_n) = \Big(\sum_{k=1}^{n} A_{1,k}x_k,\ \dots,\ \sum_{k=1}^{n} A_{n,k}x_k\Big),$$

which is linear because each coordinate is (Check!).
(i) Since $x_1, \dots, x_n$ solves the homogeneous system exactly when $(x_1, \dots, x_n) \in \operatorname{null} T$, statement (a) says $\operatorname{null} T = \{0\}$, that is, T is injective (by 3.15).
(ii) Since the system with right side $c_1, \dots, c_n$ is solvable exactly when $(c_1, \dots, c_n) \in \operatorname{range} T$, statement (b) says $\operatorname{range} T = \mathbf{F}^n$, that is, T is surjective (3.19).
Because the equations and the variables are equal in number, T is an operator on the finite-dimensional space $\mathbf{F}^n$, so 3.65 applies and gives that T is injective if and only if T is surjective. Hence (a) holds if and only if (b) holds.

Problem 3D.22. Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Prove that

$$\mathcal{M}\big(T, (v_1, \dots, v_n)\big) \text{ is invertible} \iff T \text{ is invertible.}$$

Solution. Every matrix below is taken with respect to $v_1, \dots, v_n$ in both roles, so $\mathcal{M}(I) = I$ (each $Iv_k = v_k$, by 3.79) and $\mathcal{M}(SR) = \mathcal{M}(S)\mathcal{M}(R)$ for $S, R \in \mathcal{L}(V)$ (by 3.81, with this one basis in all three roles). Write $A = \mathcal{M}(T)$; note $\dim V = n$.
($\Longrightarrow$) If T is invertible then $T^{-1} \in \mathcal{L}(V)$ and

$$A\,\mathcal{M}(T^{-1}) = \mathcal{M}(TT^{-1}) = \mathcal{M}(I) = I,$$
$$\mathcal{M}(T^{-1})\,A = \mathcal{M}(T^{-1}T) = \mathcal{M}(I) = I,$$

so A is invertible with $A^{-1} = \mathcal{M}(T^{-1})$ (3.80).
($\Longleftarrow$) If A is invertible, then because $\mathcal{M}\colon \mathcal{L}(V) \to \mathbf{F}^{n,n}$ is an isomorphism (by 3.71, with $v_1, \dots, v_n$ as basis of domain and target) there is $S \in \mathcal{L}(V)$ with $\mathcal{M}(S) = A^{-1}$, whence

$$\mathcal{M}(ST) = \mathcal{M}(S)\,\mathcal{M}(T) = A^{-1}A = I = \mathcal{M}(I).$$

Injectivity of $\mathcal{M}$ gives $ST = I$, and then $TS = I$ by 3.68 (V finite-dimensional). Hence T is invertible.

Problem 3D.23. Suppose that $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Let $T \in \mathcal{L}(V)$ be such that $Tv_k = u_k$ for each $k = 1, \dots, n$. Prove that

$$\mathcal{M}\big(T, (v_1, \dots, v_n)\big) = \mathcal{M}\big(I, (u_1, \dots, u_n), (v_1, \dots, v_n)\big).$$

Solution. The two matrices have the same columns, because $Tv_k = u_k$. Write $A = \mathcal{M}(T, (v_1, \dots, v_n))$ and $C = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$; by the definition 3.31 of the matrix of a linear map, their entries are the scalars with

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n,$$
$$u_k = Iu_k = C_{1,k}v_1 + \cdots + C_{n,k}v_n,$$

for each $k = 1, \dots, n$. Fix k; since $Tv_k = u_k$ and a vector has only one representation as a linear combination of the basis $v_1, \dots, v_n$ (by 2.28), we get $A_{j,k} = C_{j,k}$ for every j. Hence $A = C$.

Problem 3D.24. Suppose A and B are square matrices of the same size and $AB = I$. Prove that $BA = I$.

Solution. Transfer to operators on $\mathbf{F}^n$, where 3.68 applies. Let n be the common size of A and B, and take every matrix with respect to the standard basis $e_1, \dots, e_n$ of $\mathbf{F}^n$ in both roles, so that $\mathcal{M}(I) = I$ (each $Ie_k = e_k$, by 3.79) and $\mathcal{M}(RQ) = \mathcal{M}(R)\mathcal{M}(Q)$ for operators R, Q (by 3.81, with this one basis in all three roles). Since $\mathcal{M}\colon \mathcal{L}(\mathbf{F}^n) \to \mathbf{F}^{n,n}$ is an isomorphism (by 3.71, with $e_1, \dots, e_n$ as basis of domain and target), choose $S, T \in \mathcal{L}(\mathbf{F}^n)$ with $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$. Then

$$\mathcal{M}(ST) = \mathcal{M}(S)\,\mathcal{M}(T) = AB = I = \mathcal{M}(I),$$

so $ST = I$ by injectivity of $\mathcal{M}$. As $\mathbf{F}^n$ is finite-dimensional, 3.68 gives $TS = I$, and hence

$$BA = \mathcal{M}(T)\,\mathcal{M}(S) = \mathcal{M}(TS) = \mathcal{M}(I) = I.$$

3.5 Exercises 3E

Problem 3E.1. Suppose T is a function from V to W. The graph of T is the subset of $V \times W$ defined by

$$\text{graph of } T = \{(v, Tv) \in V \times W : v \in V\}.$$

Prove that T is a linear map if and only if the graph of T is a subspace of $V \times W$.
[Formally, a function T from V to W is a subset T of $V \times W$ such that for each $v \in V$, there exists exactly one element $(v, w) \in T$. In other words, formally a function is what is called above its graph. We do not usually think of functions in this formal manner. However, if we do become formal, then this exercise could be rephrased as follows: Prove that a function T from V to W is a linear map if and only if T is a subspace of $V \times W$.]

Solution. Write $G = \{(v, Tv) : v \in V\}$ for the graph. Because T is a function, for $v \in V$ and $w \in W$

$$(v, w) \in G \iff w = Tv$$

(an element (u, Tu) of G equals (v, w) only if $u = v$, forcing $w = Tv$); both directions below use this. By 1.34, G is a subspace exactly when $(0, 0) \in G$ and G is closed under addition and scalar multiplication.
($\Longrightarrow$) If T is linear then $T0 = 0$ (by 3.10), so $(0, 0) = (0, T0) \in G$, and additivity and homogeneity of T give

$$(u, Tu) + (v, Tv) = (u + v, Tu + Tv) = \big(u + v, T(u + v)\big) \in G,$$
$$\lambda(v, Tv) = (\lambda v, \lambda Tv) = \big(\lambda v, T(\lambda v)\big) \in G.$$

Hence G is a subspace of $V \times W$.
($\Longleftarrow$) If G is a subspace then for $u, v \in V$ and $\lambda \in \mathbf{F}$, closure under the two operations gives

$$(u + v,\ Tu + Tv) \in G \qquad \text{and} \qquad (\lambda v,\ \lambda Tv) \in G,$$

which by the equivalence above say $Tu + Tv = T(u + v)$ and $\lambda Tv = T(\lambda v)$. So T is linear (3.1).

Problem 3E.2. Suppose that $V_1, \ldots, V_m$ are vector spaces such that $V_1 \times \cdots \times V_m$ is finite-dimensional. Prove that V_k is finite-dimensional for each $k = 1, \ldots, m$.

Solution. Apply 3.21 to the k-th coordinate projection

$$P_k : V_1 \times \cdots \times V_m \to V_k, \qquad P_k(v_1, \ldots, v_m) = v_k,$$

which is linear because the operations on the product are slotwise (3.87) (Check!) and surjective because $P_k(0, \ldots, 0, v, 0, \ldots, 0) = v$ for each $v \in V_k$ (with v in slot k). Its domain is finite-dimensional by hypothesis, so 3.21 gives that range $P_k = V_k$ is finite-dimensional. As k was arbitrary, each V_k is finite-dimensional.

Problem 3E.3. Suppose $V_1, \ldots, V_m$ are vector spaces. Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces.
[There is no assumption in the exercise above or in the two following exercises that the vector spaces are finite-dimensional.]

Solution. Restriction to the slots is an isomorphism. Write $\mathcal{V} = V_1 \times \cdots \times V_m$ and let $\iota_k : V_k \to \mathcal{V}$ be the k-th inclusion $\iota_k(v) = (0, \ldots, 0, v, 0, \ldots, 0)$, which is linear because the operations on $\mathcal{V}$ are slotwise

(3.87) and satisfies

$$(v_1, \ldots, v_m) = \iota_1(v_1) + \cdots + \iota_m(v_m).$$

Define $\Phi(T) = (T \circ \iota_1, \ldots, T \circ \iota_m)$ for $T \in \mathcal{L}(\mathcal{V}, W)$; each $T \circ \iota_k$ lies in $\mathcal{L}(V_k, W)$ by 3.7, so Φ maps $\mathcal{L}(\mathcal{V}, W)$ into $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$, and Φ is linear because the operations on both sides are computed pointwise and slotwise (3.5, 3.87) (Check!).
Φ is injective: if $T \circ \iota_k = 0$ for every k, then the display above and linearity of T give

$$T(v_1, \ldots, v_m) = (T \circ \iota_1)(v_1) + \cdots + (T \circ \iota_m)(v_m) = 0,$$

so $T = 0$ and $\operatorname{null} \Phi = \{0\}$ (by 3.15).
Φ is surjective: given $(T_1, \ldots, T_m)$, the map $T(v_1, \ldots, v_m) = T_1 v_1 + \cdots + T_m v_m$ is linear by additivity and homogeneity of each T_k together with slotwise operations (Check!), and

$$(T \circ \iota_k)(v) = T_1 0 + \cdots + T_k v + \cdots + T_m 0 = T_k v$$

since $T_j 0 = 0$ (3.10); thus $\Phi(T) = (T_1, \ldots, T_m)$.
Hence Φ is invertible (by 3.63), so the two spaces are isomorphic (3.69).

Problem 3E.4. Suppose $W_1, \ldots, W_m$ are vector spaces. Prove that $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ and $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ are isomorphic vector spaces.

Solution. Recording the component maps is an isomorphism. Write $\mathcal{W} = W_1 \times \cdots \times W_m$ and let $P_k : \mathcal{W} \to W_k$ be the k-th coordinate projection, linear as in Exercise 2 of this section, so that

$$w = \big(P_1 w, \ldots, P_m w\big) \qquad \text{for all } w \in \mathcal{W}.$$

Define $\Psi(T) = (P_1 \circ T, \ldots, P_m \circ T)$ for $T \in \mathcal{L}(V, \mathcal{W})$; each $P_k \circ T$ lies in $\mathcal{L}(V, W_k)$ by 3.7, so Ψ maps $\mathcal{L}(V, \mathcal{W})$ into $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$, and Ψ is linear because the operations on both sides are pointwise and slotwise (3.5, 3.87) and each P_k is linear (Check!).
Ψ is injective: if $P_k(Tv) = 0$ for every k and every $v \in V$, then the display above applied to $w = Tv$ gives

$$Tv = \big(P_1(Tv), \ldots, P_m(Tv)\big) = (0, \ldots, 0),$$

so $T = 0$ and $\operatorname{null} \Psi = \{0\}$ (by 3.15).
Ψ is surjective: given $(T_1, \ldots, T_m)$, the map $Tv = (T_1 v, \ldots, T_m v)$ is linear by additivity and homogeneity of each T_k together with slotwise operations in $\mathcal{W}$ (Check!), and $(P_k \circ T)v = T_k v$ for all k and v, so $\Psi(T) = (T_1, \ldots, T_m)$.
Hence Ψ is invertible (by 3.63), so the two spaces are isomorphic (3.69).

Problem 3E.5. For m a positive integer, define V^m by

$$V^m = \underbrace{V \times \cdots \times V}_{m \text{ times}}.$$

Prove that V^m and $\mathcal{L}(\mathbf{F}^m, V)$ are isomorphic vector spaces.

Solution. Evaluation on the standard basis is an isomorphism: with $e_1, \ldots, e_m$ the standard basis of $\mathbf{F}^m$, define

$$\Phi : \mathcal{L}(\mathbf{F}^m, V) \to V^m, \qquad \Phi(T) = (Te_1, \ldots, Te_m),$$

linear because the operations on $\mathcal{L}(\mathbf{F}^m, V)$ are pointwise (3.5) and those on V^m slotwise (3.87) (Check!).
Φ is injective: if $Te_k = 0$ for each k, then for every $(x_1, \ldots, x_m) \in \mathbf{F}^m$,

$$T(x_1, \ldots, x_m) = x_1 Te_1 + \cdots + x_m Te_m = 0,$$

so $T = 0$ and $\operatorname{null} \Phi = \{0\}$ (by 3.15).

Φ is surjective: given $(v_1, \ldots, v_m) \in V^m$, the linear map lemma 3.4 applied to the basis $e_1, \ldots, e_m$ provides $T \in \mathcal{L}(\mathbf{F}^m, V)$ with $Te_k = v_k$ for each k, so $\Phi(T) = (v_1, \ldots, v_m)$.
Hence Φ is invertible (by 3.63), so $\mathcal{L}(\mathbf{F}^m, V)$ and V^m are isomorphic (3.69).

Problem 3E.6. Suppose that v, x are vectors in V and that U, W are subspaces of V such that $v + U = x + W$. Prove that $U = W$.

Solution. We prove $U \subseteq W$; symmetry then finishes. Since $0 \in U$, we have $v \in v + U = x + W$, so $v = x + w_0$ for some $w_0 \in W$. Given $u \in U$, also $v + u \in v + U = x + W$, say $v + u = x + w$ with $w \in W$, whence

$$u = (x + w) - v = (x + w) - (x + w_0) = w - w_0 \in W,$$

as W is closed under additive inverses and addition. Thus $U \subseteq W$. Interchanging the pairs (v, U) and (x, W), which the symmetric hypothesis $v + U = x + W$ permits, gives $W \subseteq U$. Hence $U = W$.

Problem 3E.7. Let $U = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$. Suppose $A \subseteq \mathbf{R}^3$. Prove that A is a translate of U if and only if there exists $c \in \mathbf{R}$ such that

$$A = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = c\}.$$

Solution. The translates of U are exactly the level sets of $\varphi(x, y, z) = 2x + 3y + 5z$. This $\varphi : \mathbf{R}^3 \to \mathbf{R}$ is linear (Check!), so $U = \operatorname{null} \varphi$ is a subspace of $\mathbf{R}^3$ (by 3.13); write $A_c = \{p \in \mathbf{R}^3 : \varphi(p) = c\}$, which is the set displayed in the exercise. Since $p \in v + U$ means $p - v \in U = \operatorname{null} \varphi$, linearity of φ gives, for $v, p \in \mathbf{R}^3$,

$$p \in v + U \iff \varphi(p - v) = 0 \iff \varphi(p) = \varphi(v),$$

that is,

$$v + U = A_{\varphi(v)} \qquad \text{for every } v \in \mathbf{R}^3.$$

($\Longrightarrow$) If A is a translate of U, say $A = v + U$ (3.97), then $A = A_c$ with $c = \varphi(v)$.
($\Longleftarrow$) If $A = A_c$, then $v = \left(\frac{c}{2}, 0, 0\right)$ has $\varphi(v) = c$, so $v + U = A_c = A$ and A is a translate of U.

Problem 3E.8. (a) Suppose $T \in \mathcal{L}(V, W)$ and $c \in W$. Prove that $\{x \in V : Tx = c\}$ is either the empty set or is a translate of $\operatorname{null} T$.
(b) Explain why the set of solutions to a system of linear equations such as 3.27 is either the empty set or is a translate of some subspace of $\mathbf{F}^n$.

Solution. (a) If $A = \{x \in V : Tx = c\}$ is nonempty, fix $x_0 \in A$; then $A = x_0 + \operatorname{null} T$, since linearity of T and $Tx_0 = c$ give, for $x \in V$,

$$\begin{aligned} x \in A &\iff T(x - x_0) = Tx - c = 0 \iff x - x_0 \in \operatorname{null} T \\ &\iff x \in x_0 + \operatorname{null} T. \end{aligned}$$

As $\operatorname{null} T$ is a subspace of V (by 3.13), A is then a translate of it (3.97).
(b) The system 3.27 says $Tx = c$, where $c = (c_1, \ldots, c_m)$ and $T\colon \mathbf{F}^n \to \mathbf{F}^m$ is the linear map of 3.25,

$$T(x_1, \ldots, x_n) = \Big(\sum_{k=1}^{n} A_{1,k} x_k, \ \ldots, \ \sum_{k=1}^{n} A_{m,k} x_k \Big).$$

So by (a) its solution set is either empty or a translate of $\operatorname{null} T$, a subspace of $\mathbf{F}^n$ (by 3.13).

Problem 3E.9. Prove that a nonempty subset A of V is a translate of some subspace of V if and only if $\lambda v + (1 - \lambda) w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$.

Solution. ($\Longrightarrow$) If $A = x + U$ with U a subspace, write $v = x + u_1$ and $w = x + u_2$ with $u_1, u_2 \in U$; then for $\lambda \in \mathbf{F}$,

$$\begin{aligned}\lambda v + (1-\lambda)w &= \lambda(x+u_1) + (1-\lambda)(x+u_2)\\ &= x + \big(\lambda u_1 + (1-\lambda)u_2\big) \in x + U = A,\end{aligned}$$

since U is closed under scalar multiplication and addition.
($\Longleftarrow$) Assume $(*)$: $\lambda v + (1-\lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$. Fix $x \in A$ (possible as $A \neq \varnothing$) and set $U = A - x = \{a - x : a \in A\}$, so that $A = x + U$; we check the three conditions of 1.34 for U.
(i) Additive identity: $0 = x - x \in U$, since $x \in A$.
(ii) Scalar multiples: for $u = a - x \in U$ and $\lambda \in \mathbf{F}$, condition $(*)$ with $v = a$, $w = x$ gives

$$x + \lambda u = \lambda a + (1-\lambda)x \in A, \qquad \text{so} \qquad \lambda u \in A - x = U.$$

(iii) Sums: for $u_1 = a - x$ and $u_2 = b - x$ in U, condition $(*)$ with $\lambda = \frac{1}{2}$ (a scalar, as $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$) gives $\frac{1}{2}a + \frac{1}{2}b \in A$, hence

$$\tfrac{1}{2}u_1 + \tfrac{1}{2}u_2 = \big(\tfrac{1}{2}a + \tfrac{1}{2}b\big) - x \in U,$$

and applying (ii) with the scalar 2 yields $u_1 + u_2 \in U$.
Thus U is a subspace and $A = x + U$ is a translate of it.

Problem 3E.10. Suppose $A_1 = v + U_1$ and $A_2 = w + U_2$ for some $v, w \in V$ and some subspaces U_1, U_2 of V. Prove that the intersection $A_1 \cap A_2$ is either a translate of some subspace of V or is the empty set.

Solution. If $A_1 \cap A_2 \neq \varnothing$, fix $x \in A_1 \cap A_2$; then $A_1 \cap A_2 = x + (U_1 \cap U_2)$, a translate of the subspace $U_1 \cap U_2$ of V (Exercise 1C.10). Indeed $x - v \in U_1$ and $x - w \in U_2$, so 3.101 gives $A_1 = x + U_1$ and $A_2 = x + U_2$, whence for $y \in V$,

$$\begin{aligned} y \in A_1 \cap A_2 &\iff y - x \in U_1 \text{ and } y - x \in U_2\\ &\iff y - x \in U_1 \cap U_2 \iff y \in x + (U_1 \cap U_2).\end{aligned}$$

Otherwise $A_1 \cap A_2$ is the empty set.

Problem 3E.11. Suppose $U = \{(x_1, x_2, \dots) \in \mathbf{F}^\infty : x_k \neq 0 \text{ for only finitely many } k\}$.
(a) Show that U is a subspace of $\mathbf{F}^\infty$.
(b) Prove that $\mathbf{F}^\infty/U$ is infinite-dimensional.

Solution. Write $\operatorname{supp} x = \{k \in \mathbf{Z}^+ : x_k \neq 0\}$, so that U consists of the $x \in \mathbf{F}^\infty$ with $\operatorname{supp} x$ finite.
(a) The three conditions of 1.34 follow from

$$\operatorname{supp} 0 = \varnothing, \qquad \operatorname{supp}(\lambda x) \subseteq \operatorname{supp} x,$$

$$\operatorname{supp}(x+y) \subseteq \operatorname{supp} x \cup \operatorname{supp} y,$$

together with the finiteness of a union of two finite sets. Hence U is a subspace of $\mathbf{F}^\infty$.
(b) For each $j \in \mathbf{Z}^+$ let

$$S_j = \{2^{j-1}(2k-1) : k \in \mathbf{Z}^+\},$$

and let $v_j \in \mathbf{F}^\infty$ have k-th coordinate 1 for $k \in S_j$ and 0 otherwise. Since every $m \in \mathbf{Z}^+$ is $2^a q$ with q odd in exactly one way, m lies in S_{a+1} alone; so the S_j are pairwise disjoint, and each is infinite.
Fix $n \in \mathbf{Z}^+$; then $v_1 + U, \dots, v_n + U$ is linearly independent in $\mathbf{F}^\infty/U$. For if $a_1(v_1+U) + \dots + a_n(v_n+U) = 0 + U$, then by the definition 3.102 of the quotient operations and by 3.101,

$$x = a_1 v_1 + \dots + a_n v_n \in U,$$

while disjointness of $S_1, \dots, S_n$ makes exactly one of $(v_1)_k, \dots, (v_n)_k$ equal to 1 for $k \in S_j$, so $x_k = a_j$ there. Were some $a_j \neq 0$, the infinite set S_j would lie in $\operatorname{supp} x$, contradicting $x \in U$; hence $a_1 = \dots = a_n = 0$.

So $\mathbf{F}^\infty/U$ has linearly independent lists of every length, which a finite-dimensional space cannot (by 2.22). Hence $\mathbf{F}^\infty/U$ is infinite-dimensional.

Problem 3E.12. Suppose $v_1, \dots, v_m \in V$. Let

$$A = \{\lambda_1 v_1 + \cdots + \lambda_m v_m : \lambda_1, \dots, \lambda_m \in \mathbf{F} \text{ and } \lambda_1 + \cdots + \lambda_m = 1\}.$$

(a) Prove that A is a translate of some subspace of V.
(b) Prove that if B is a translate of some subspace of V and $\{v_1, \dots, v_m\} \subseteq B$, then $A \subseteq B$.
(c) Prove that A is a translate of some subspace of V of dimension less than m.

Solution. $A = v_1 + U$, where $U = \operatorname{span}(v_2 - v_1, \dots, v_m - v_1)$ is a subspace of V by 2.6; this single identity gives (a), and gives (c) because U is spanned by a list of length $m-1$, which contains a basis of U by 2.30, so $\dim U \le m - 1 < m$.
(a) and (c). If $a = \lambda_1 v_1 + \cdots + \lambda_m v_m$ with $\sum_j \lambda_j = 1$, then

$$\begin{aligned} a - v_1 &= \sum_{j=1}^m \lambda_j v_j - \Big(\sum_{j=1}^m \lambda_j\Big) v_1 \\ &= \sum_{j=1}^m \lambda_j (v_j - v_1) = \sum_{j=2}^m \lambda_j (v_j - v_1) \in U, \end{aligned}$$

the last step because the $j = 1$ term vanishes; hence $A \subseteq v_1 + U$. Conversely, if $u = c_2(v_2 - v_1) + \cdots + c_m(v_m - v_1)$, then

$$v_1 + u = \Big(1 - \sum_{j=2}^m c_j\Big) v_1 + c_2 v_2 + \cdots + c_m v_m,$$

whose coefficients sum to 1, so $v_1 + u \in A$. Thus $A = v_1 + U$.
(b) Write $B = x + W$ with $x \in V$ and W a subspace of V; since each $v_j \in B$, we have $v_j - x \in W$. For $a = \lambda_1 v_1 + \cdots + \lambda_m v_m \in A$ with $\sum_j \lambda_j = 1$,

$$a - x = \sum_{j=1}^m \lambda_j v_j - \Big(\sum_{j=1}^m \lambda_j\Big) x = \sum_{j=1}^m \lambda_j (v_j - x) \in W,$$

so $a \in x + W = B$. Hence $A \subseteq B$.

Problem 3E.13. Suppose U is a subspace of V such that V/U is finite-dimensional. Prove that V is isomorphic to $U \times (V/U)$.

Solution. The map $S(u, x) = u + \Psi x$ is an isomorphism from $U \times (V/U)$ onto V, where Ψ is defined as follows. Let $v_1 + U, \dots, v_n + U$ be a basis of V/U (its vectors have this form by 3.99), and let $\Psi \in \mathcal{L}(V/U, V)$ be the linear map with $\Psi(v_j + U) = v_j$ for each j, which exists by the linear map lemma 3.4 applied to that basis.
Since the operations on $U \times (V/U)$ are coordinatewise (3.87) and Ψ is linear, S is linear (Check!).
S is surjective: given $v \in V$, write $v + U = a_1(v_1 + U) + \cdots + a_n(v_n + U)$, which equals $(a_1 v_1 + \cdots + a_n v_n) + U$ by 3.102, so $u := v - (a_1 v_1 + \cdots + a_n v_n) \in U$ by 3.101; with $x = v + U$, linearity of Ψ gives $\Psi x = a_1 v_1 + \cdots + a_n v_n$ and hence

$$S(u, x) = u + \Psi x = v.$$

S is injective: if $u + \Psi x = 0$, write $x = a_1(v_1 + U) + \cdots + a_n(v_n + U)$, so $\Psi x = a_1 v_1 + \cdots + a_n v_n = -u \in U$; then 3.101 and 3.102 give

$$x = (a_1 v_1 + \cdots + a_n v_n) + U = 0 + U,$$

whence $\Psi x = 0$ and $u = 0$. So $\operatorname{null} S = \{0\}$ and S is injective by 3.15.

Being linear, injective, and surjective, S is invertible by 3.63, hence an isomorphism by 3.69.

Problem 3E.14. Suppose U and W are subspaces of V and $V = U \oplus W$. Suppose $w_1, \dots, w_m$ is a basis of W. Prove that $w_1 + U, \dots, w_m + U$ is a basis of V/U.

Solution. The list spans V/U and is linearly independent, hence is a basis.
Spanning. Every element of V/U is $v+U$ for some $v \in V$ (3.99); as $V = U + W$, write $v = u + w$ with $u \in U$ and $w = a_1w_1 + \cdots + a_mw_m \in W$. Then $v - (a_1w_1 + \cdots + a_mw_m) = u \in U$, so 3.101 and 3.102 give

$$v + U = (a_1w_1 + \cdots + a_mw_m) + U = a_1(w_1 + U) + \cdots + a_m(w_m + U).$$

Linear independence. Suppose $a_1(w_1 + U) + \cdots + a_m(w_m + U) = 0 + U$. By 3.102 the left side is $(a_1w_1 + \cdots + a_mw_m) + U$, so 3.101 gives

$$a_1w_1 + \cdots + a_mw_m \in U \cap W = \{0\},$$

the last equality by 1.46 since $V = U \oplus W$ (the vector lies in W because W is a subspace containing each w_k). Linear independence of the basis $w_1, \dots, w_m$ of W now forces $a_1 = \cdots = a_m = 0$.

Problem 3E.15. Suppose U is a subspace of V and $v_1 + U, \dots, v_m + U$ is a basis of V/U and $u_1, \dots, u_n$ is a basis of U. Prove that $v_1, \dots, v_m, u_1, \dots, u_n$ is a basis of V.

Solution. The list spans V and is linearly independent, hence is a basis of V (2.26). Let $\pi\colon V \to V/U$ be the quotient map $\pi(v) = v + U$ (3.104); by 3.101, $\operatorname{null} \pi = U$.
Spanning. Given $v \in V$, write $v + U = a_1(v_1 + U) + \cdots + a_m(v_m + U)$, which by 3.102 equals $(a_1v_1 + \cdots + a_mv_m) + U$, so 3.101 gives $v - (a_1v_1 + \cdots + a_mv_m) \in U$; expanding that vector in the basis $u_1, \dots, u_n$ of U yields

$$v = a_1v_1 + \cdots + a_mv_m + b_1u_1 + \cdots + b_nu_n.$$

Linear independence. Suppose $a_1v_1 + \cdots + a_mv_m + b_1u_1 + \cdots + b_nu_n = 0$. Applying π and using $u_j \in \operatorname{null} \pi$ gives

$$a_1(v_1 + U) + \cdots + a_m(v_m + U) = 0 + U,$$

so $a_1 = \cdots = a_m = 0$ by linear independence of the basis $v_1 + U, \dots, v_m + U$. The displayed equation then reduces to $b_1u_1 + \cdots + b_nu_n = 0$, and linear independence of $u_1, \dots, u_n$ gives $b_1 = \cdots = b_n = 0$.

Problem 3E.16. Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$ and $\varphi \neq 0$. Prove that $\dim V/(\operatorname{null} \varphi) = 1$.

Solution. Pick $v \in V$ with $\varphi(v) \neq 0$, which exists because $\varphi \neq 0$; then the one-vector list $v + \operatorname{null} \varphi$ is a basis of $V/(\operatorname{null} \varphi)$, so that space has dimension 1. (Note V is not assumed finite-dimensional, so 3.105 is unavailable.)
It spans: for every $x \in V$ we have $\varphi\big(x - \frac{\varphi(x)}{\varphi(v)}v\big) = \varphi(x) - \varphi(x) = 0$, so by 3.101,

$$x + \operatorname{null} \varphi = \frac{\varphi(x)}{\varphi(v)}\,\big(v + \operatorname{null} \varphi\big).$$

It is linearly independent: $\lambda(v + \operatorname{null} \varphi) = 0 + \operatorname{null} \varphi$ gives $\lambda v \in \operatorname{null} \varphi$, so $\lambda\varphi(v) = 0$ and hence $\lambda = 0$.
Method (2): since $\varphi\big(\frac{\lambda}{\varphi(v)}v\big) = \lambda$ for every $\lambda \in \mathbf{F}$, we have $\operatorname{range} \varphi = \mathbf{F}$, so 3.107(d) makes $\widetilde{\varphi}$ of 3.106 an isomorphism from $V/(\operatorname{null} \varphi)$ onto $\mathbf{F}$. Its inverse $\widetilde{\varphi}^{-1}$ is an isomorphism too (3.69, 3.63), and an isomorphism carries a basis to a basis; transporting the basis 1 of $\mathbf{F}$ (the direction that matters, since a basis of the quotient is what we lack) gives the one-vector basis $\widetilde{\varphi}^{-1}(1)$.

Problem 3E.17. Suppose U is a subspace of V such that $\dim V/U = 1$. Prove that there exists $\varphi \in \mathcal{L}(V, \mathbf{F})$ such that $\operatorname{null} \varphi = U$.

Solution. Take $\varphi = \alpha \circ \pi$, where $\pi \colon V \to V/U$ is the quotient map $\pi(v) = v + U$ (3.104) and $\alpha \in (V/U)'$ is the linear functional determined by $\alpha(v + U) = 1$ for a one-vector basis $v + U$ of V/U (such a basis exists because $\dim V/U = 1$, and α exists by the linear map lemma 3.4). Then $\varphi \in \mathcal{L}(V, \mathbf{F})$, since a composition of linear maps is linear (3.7).
By 3.101, $\operatorname{null} \pi = U$. Also α is injective: $\alpha\big(\lambda(v + U)\big) = \lambda$, so $\operatorname{null} \alpha = \{0\}$ and 3.15 applies. Hence

$$\operatorname{null} \varphi = \{x \in V : \alpha(\pi x) = 0\} = \{x \in V : \pi x = 0\} = \operatorname{null} \pi = U.$$

Problem 3E.18. Suppose that U is a subspace of V such that V/U is finite-dimensional.
(a) Show that if W is a finite-dimensional subspace of V and $V = U + W$, then $\dim W \geq \dim V/U$.
(b) Prove that there exists a finite-dimensional subspace W of V such that $\dim W = \dim V/U$ and $V = U \oplus W$.

Solution. Throughout, $\pi \colon V \to V/U$ is the quotient map $\pi(v) = v + U$ (3.104), which is linear with $\operatorname{null} \pi = U$ by 3.101.
(a) The restriction $\pi|_W \in \mathcal{L}(W, V/U)$ is surjective: given $v + U \in V/U$, write $v = u + w$ with $u \in U$ and $w \in W$ (possible as $V = U + W$), so $v - w \in U$ and hence $v + U = w + U = \pi|_W(w)$ by 3.101. Since W is finite-dimensional, the fundamental theorem of linear maps 3.21 gives

$$\dim W = \dim(U \cap W) + \dim V/U \geq \dim V/U,$$

using $\operatorname{null} \pi|_W = W \cap \operatorname{null} \pi = U \cap W$.
(b) Take $W = \operatorname{span}(v_1, \dots, v_n)$, where $v_1 + U, \dots, v_n + U$ is a basis of V/U and $n = \dim V/U$ (such a basis exists by 2.31, and its vectors have this form because π is surjective).
The list $v_1, \dots, v_n$ is linearly independent, since applying π to $a_1 v_1 + \cdots + a_n v_n = 0$ gives $a_1(v_1 + U) + \cdots + a_n(v_n + U) = 0 + U$ and hence $a_1 = \cdots = a_n = 0$; so $\dim W = n = \dim V/U$.
$V = U + W$: for $v \in V$, expanding $v + U$ in the basis and using 3.102 gives $v + U = (a_1 v_1 + \cdots + a_n v_n) + U$, so by 3.101

$$v = \big(v - (a_1 v_1 + \cdots + a_n v_n)\big) + (a_1 v_1 + \cdots + a_n v_n) \in U + W.$$

$U \cap W = \{0\}$: if $w = a_1 v_1 + \cdots + a_n v_n$ lies in U, then $\pi(w) = a_1(v_1 + U) + \cdots + a_n(v_n + U) = 0 + U$, so every $a_j = 0$ and $w = 0$.
By 1.46 the sum is direct, so $V = U \oplus W$ with $\dim W = \dim V/U$.

Problem 3E.19. Suppose $T \in \mathcal{L}(V, W)$ and U is a subspace of V. Let π denote the quotient map from V onto V/U. Prove that there exists $S \in \mathcal{L}(V/U, W)$ such that $T = S \circ \pi$ if and only if $U \subseteq \operatorname{null} T$.

Solution. If such an S exists, then each $u \in U$ has $\pi(u) = 0 + U$ by 3.101, so

$$Tu = S(\pi(u)) = S(0) = 0$$

by 3.10; hence $U \subseteq \operatorname{null} T$.
Conversely, suppose $U \subseteq \operatorname{null} T$ and set $S(v + U) = Tv$. This is well defined (the point of the hypothesis): if $v + U = v' + U$, then $v - v' \in U \subseteq \operatorname{null} T$ by 3.101, so $Tv = Tv'$. It is linear because addition and scalar multiplication on V/U are given by 3.102 and T is linear:

$$S\big((v + U) + \lambda(v' + U)\big) = T(v + \lambda v') = S(v + U) + \lambda S(v' + U).$$

Finally $(S \circ \pi)(v) = S(v + U) = Tv$ for every $v \in V$, so $T = S \circ \pi$.

3.6 Exercises 3F

Problem 3F.1. Explain why each linear functional is surjective or is the zero map.

Solution. A linear functional $\varphi \in V'$ that is not the zero map is surjective. Indeed, choose $u \in V$ with $a := \varphi(u) \neq 0$; since a is an invertible element of the field $\mathbf{F}$, homogeneity of φ gives, for every $\lambda \in \mathbf{F}$,

$$\varphi\left(\frac{\lambda}{a}u\right) = \frac{\lambda}{a}\varphi(u) = \lambda,$$

so $\operatorname{range}\varphi = \mathbf{F}$.
Method (2): $\operatorname{range}\varphi$ is a subspace of $\mathbf{F}$ by 3.18, and $\dim \mathbf{F} = 1$ (the list 1 is a basis), so $\dim \operatorname{range}\varphi \leq 1$ by 2.37. If that dimension is 0 then φ is the zero map; if it is 1 then $\operatorname{range}\varphi = \mathbf{F}$ by 2.39.

Problem 3F.2. Give three distinct examples of linear functionals on $\mathbf{R}^{[0,1]}$.

Solution. Take

$$\varphi_1(f) = f(0), \qquad \varphi_2(f) = f(1), \qquad \varphi_3(f) = f(0) + f(1).$$

Each evaluation $f \mapsto f(t)$ is a linear functional on $\mathbf{R}^{[0,1]}$, because the operations there are pointwise: $(f + \lambda g)(t) = f(t) + \lambda g(t)$. This gives φ_1 (take $t = 0$) and φ_2 (take $t = 1$), and $\varphi_3 = \varphi_1 + \varphi_2$ is linear as a sum of linear maps (3.6).
They are pairwise distinct: let f equal 1 at 0 and 0 on $(0, 1]$, and let g equal 1 at 1 and 0 on $[0, 1)$. Then

$$\big(\varphi_1(f), \varphi_2(f), \varphi_3(f)\big) = (1, 0, 1), \qquad \big(\varphi_1(g), \varphi_2(g), \varphi_3(g)\big) = (0, 1, 1),$$

so f separates φ_1 from φ_2 and φ_2 from φ_3, while g separates φ_1 from φ_3.

Problem 3F.3. Suppose V is finite-dimensional and $v \in V$ with $v \neq 0$. Prove that there exists $\varphi \in V'$ such that $\varphi(v) = 1$.

Solution. Take φ to be the first functional in the dual basis of a basis of V whose first vector is v. In detail: the one-vector list v is linearly independent (if $av = 0$ with $a \neq 0$, then $v = a^{-1}(av) = 0$, contrary to $v \neq 0$), so since V is finite-dimensional it extends by 2.32 to a basis $v_1, \dots, v_n$ of V with $v_1 = v$. By the linear map lemma 3.4 there is $\varphi \in \mathcal{L}(V, \mathbf{F}) = V'$ with

$$\varphi(v_1) = 1 \qquad \text{and} \qquad \varphi(v_j) = 0 \quad \text{for } j = 2, \dots, n,$$

and then $\varphi(v) = \varphi(v_1) = 1$.

Problem 3F.4. Suppose V is finite-dimensional and U is a subspace of V such that $U \neq V$. Prove that there exists $\varphi \in V'$ such that $\varphi(u) = 0$ for every $u \in U$ but $\varphi \neq 0$.

Solution. Extend a basis $u_1, \dots, u_m$ of U to a basis $u_1, \dots, u_m, v_{m+1}, \dots, v_n$ of V (by 2.32), and let $\varphi \in V'$ be the linear functional supplied by the linear map lemma 3.4 with

$$\varphi(u_j) = 0 \ (1 \leq j \leq m), \quad \varphi(v_{m+1}) = 1, \quad \varphi(v_k) = 0 \ (m + 2 \leq k \leq n).$$

The vector v_{m+1} really is present: U is finite-dimensional by 2.25 with $\dim U \leq \dim V$ by 2.37, and $\dim U = \dim V$ would give $U = V$ by 2.39, so $m = \dim U < \dim V = n$.
Then φ vanishes on U, since every $u \in U$ is $a_1 u_1 + \cdots + a_m u_m$ and linearity gives $\varphi(u) = 0$; and $\varphi \neq 0$ because $\varphi(v_{m+1}) = 1$.

Problem 3F.5. Suppose $T \in \mathcal{L}(V, W)$ and $w_1, \ldots, w_m$ is a basis of $\operatorname{range} T$. Hence for each $v \in V$, there exist unique numbers $\varphi_1(v), \ldots, \varphi_m(v)$ such that

$$Tv = \varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m,$$

thus defining functions $\varphi_1, \ldots, \varphi_m$ from V to $\mathbf{F}$. Show that each of the functions $\varphi_1, \ldots, \varphi_m$ is a linear functional on V.

Solution. Each $\varphi_j = \psi_j \circ S$ is a composition of linear maps and hence a linear functional on V by 3.7. Here $S \in \mathcal{L}(V, \operatorname{range} T)$ is T with its target narrowed to $\operatorname{range} T$, and $\psi_1, \ldots, \psi_m \in (\operatorname{range} T)'$ is the dual basis of $w_1, \ldots, w_m$; by 3.114 the scalar $\psi_j(w)$ is the j-th coefficient of w in that basis, which for $w = Tv$ is exactly $\varphi_j(v)$ by the defining equation.
Method (2): compare coefficients. Additivity of T gives

$$\begin{aligned}\varphi_1(u+v)w_1 + \cdots + \varphi_m(u+v)w_m &= T(u+v) = Tu + Tv\\ &= \big(\varphi_1(u) + \varphi_1(v)\big)w_1 + \cdots + \big(\varphi_m(u) + \varphi_m(v)\big)w_m,\end{aligned}$$

and the uniqueness of coefficients relative to the basis $w_1, \ldots, w_m$ (2.28) forces $\varphi_j(u+v) = \varphi_j(u) + \varphi_j(v)$. Homogeneity is the same computation started from $T(\lambda v) = \lambda Tv$.

Problem 3F.6. Suppose $\varphi, \beta \in V'$. Prove that $\operatorname{null} \varphi \subseteq \operatorname{null} \beta$ if and only if there exists $c \in \mathbf{F}$ such that $\beta = c\varphi$.

Solution. If $\beta = c\varphi$ and $\varphi(v) = 0$, then $\beta(v) = c\,\varphi(v) = 0$, so $\operatorname{null} \varphi \subseteq \operatorname{null} \beta$.
Conversely, suppose $\operatorname{null} \varphi \subseteq \operatorname{null} \beta$.
(i) $\varphi = 0$. Then $V = \operatorname{null} \varphi \subseteq \operatorname{null} \beta$, so $\beta = 0 = 0 \cdot \varphi$.
(ii) $\varphi \neq 0$. By Exercise 3F.1 a nonzero linear functional is surjective, so choose $u \in V$ with $\varphi(u) = 1$ and put $c = \beta(u)$. For $v \in V$, the vector $w = v - \varphi(v)u$ satisfies $\varphi(w) = \varphi(v) - \varphi(v) = 0$, so $w \in \operatorname{null} \varphi \subseteq \operatorname{null} \beta$; applying β to $v = w + \varphi(v)u$ then gives

$$\beta(v) = \beta(w) + \varphi(v)\,\beta(u) = c\,\varphi(v) = (c\varphi)(v).$$

Hence $\beta = c\varphi$.

Problem 3F.7. Suppose that $V_1, \ldots, V_m$ are vector spaces. Prove that $(V_1 \times \cdots \times V_m)'$ and ${V_1}' \times \cdots \times {V_m}'$ are isomorphic vector spaces.

Solution. Write $V = V_1 \times \cdots \times V_m$ and let $\iota_j \colon V_j \to V$ be the inclusion $\iota_j(v) = (0, \ldots, 0, v, 0, \ldots, 0)$ with v in slot j. Then

$$\Gamma \colon V' \to {V_1}' \times \cdots \times {V_m}', \qquad \Gamma(\varphi) = (\varphi \circ \iota_1, \ldots, \varphi \circ \iota_m),$$

is an isomorphism. Each ι_j is linear because the operations on V are coordinatewise (3.87, 3.89), so each $\varphi \circ \iota_j$ lies in V_j' by 3.7; and Γ is linear because addition and scalar multiplication in V' are pointwise (3.5) and in the product are coordinatewise (Check!).
Its two-sided inverse is

$$\Lambda(\varphi_1, \ldots, \varphi_m) \colon (v_1, \ldots, v_m) \mapsto \varphi_1(v_1) + \cdots + \varphi_m(v_m),$$

which lies in V' since the operations on V are coordinatewise and each φ_j is linear (Check!). Indeed, for $v \in V_j$, using $\varphi_k(0) = 0$ (3.10),

$$\big(\Lambda(\varphi_1, \ldots, \varphi_m) \circ \iota_j\big)(v) = \varphi_j(v),$$

so $\Gamma \circ \Lambda$ is the identity; and since $(v_1, \ldots, v_m) = \iota_1(v_1) + \cdots + \iota_m(v_m)$, linearity of φ gives

$$\big(\Lambda(\Gamma\varphi)\big)(v_1, \ldots, v_m) = \varphi(\iota_1 v_1) + \cdots + \varphi(\iota_m v_m) = \varphi(v_1, \ldots, v_m),$$

so $\Lambda \circ \Gamma$ is the identity. Hence Γ is invertible (3.59), that is, an isomorphism (3.69).

Problem 3F.8. Suppose $v_1, \ldots, v_n$ is a basis of V and $\varphi_1, \ldots, \varphi_n$ is the dual basis of V'. Define $\Gamma\colon V \to \mathbf{F}^n$ and $\Lambda\colon \mathbf{F}^n \to V$ by

$$\Gamma(v) = (\varphi_1(v), \ldots, \varphi_n(v)) \quad \text{and} \quad \Lambda(a_1, \ldots, a_n) = a_1 v_1 + \cdots + a_n v_n.$$

Explain why Γ and Λ are inverses of each other.

Solution. Both composites are identity maps, so by 3.59 each of Γ, Λ is an inverse of the other (and inverses are unique by 3.60). Both are linear: Γ because each φ_j is linear and the operations on $\mathbf{F}^n$ are coordinatewise, Λ because $\sum_k (a_k + \lambda b_k) v_k = \sum_k a_k v_k + \lambda \sum_k b_k v_k$.
For $v \in V$, the expansion 3.114 of v in the basis $v_1, \ldots, v_n$ gives

$$(\Lambda \circ \Gamma)(v) = \varphi_1(v) v_1 + \cdots + \varphi_n(v) v_n = v.$$

For $(a_1, \ldots, a_n) \in \mathbf{F}^n$, putting $v = a_1 v_1 + \cdots + a_n v_n$ and using $\varphi_j(v_k) = 1$ for $k = j$ and 0 otherwise (3.112), linearity gives $\varphi_j(v) = \sum_k a_k \varphi_j(v_k) = a_j$, so

$$(\Gamma \circ \Lambda)(a_1, \ldots, a_n) = (\varphi_1(v), \ldots, \varphi_n(v)) = (a_1, \ldots, a_n).$$

Problem 3F.9. Suppose m is a positive integer. Show that the dual basis of the basis $1, x, \ldots, x^m$ of $\mathcal{P}_m(\mathbf{R})$ is $\varphi_0, \varphi_1, \ldots, \varphi_m$, where

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!}.$$

Here $p^{(k)}$ denotes the k^{th} derivative of p, with the understanding that the 0^{th} derivative of p is p.

Solution. Each φ_k is a linear functional on $\mathcal{P}_m(\mathbf{R})$ satisfying $\varphi_k(x^j) = 1$ if $j = k$ and $\varphi_k(x^j) = 0$ if $j \neq k$, which is exactly the defining condition 3.112 for the dual basis of $1, x, \ldots, x^m$.
Linearity: k-fold differentiation, evaluation at 0, and division by the fixed nonzero scalar $k!$ are each linear, so

$$\varphi_k(p + \lambda q) = \frac{p^{(k)}(0) + \lambda\, q^{(k)}(0)}{k!} = \varphi_k(p) + \lambda \varphi_k(q).$$

Duality: repeated differentiation gives $(x^j)^{(k)} = j(j-1)\cdots(j-k+1)\, x^{j-k}$ for $k \leq j$, and $(x^j)^{(k)} = 0$ for $k > j$.
(i) $k > j$: the derivative is 0, so $\varphi_k(x^j) = 0$.
(ii) $k < j$: the derivative has the factor x^{j-k} with $j - k \geq 1$, which vanishes at 0, so $\varphi_k(x^j) = 0$.
(iii) $k = j$: the derivative is the constant $k(k-1)\cdots 1 = k!$, so $\varphi_k(x^k) = k!/k! = 1$.

Problem 3F.10. Suppose m is a positive integer.
(a) Show that $1, x-5, \ldots, (x-5)^m$ is a basis of $\mathcal{P}_m(\mathbf{R})$.
(b) What is the dual basis of the basis in (a)?

Solution. (a) Each $(x-5)^j$ has degree $j \leq m$, so the list lies in $\mathcal{P}_m(\mathbf{R})$, and it has length $m + 1 = \dim \mathcal{P}_m(\mathbf{R})$ (2.36); by 2.38 it is therefore enough to check linear independence. Suppose $a_0 + a_1(x - 5) + \cdots + a_m(x-5)^m = 0$ with not all a_j zero, and let k be the largest index with $a_k \neq 0$. The terms with $j < k$ have degree at most $k - 1$, while $(x-5)^k$ has leading coefficient 1, so the coefficient of x^k in the sum is $a_k \neq 0$ and the sum is not the zero polynomial, a contradiction.
(b) The dual basis is $\varphi_0, \ldots, \varphi_m$ with

$$\varphi_k(p) = \frac{p^{(k)}(5)}{k!},$$

that is, Exercise 3F.9 with derivatives evaluated at 5 instead of 0.

Each φ_k is linear, since differentiating k times, evaluating at 5, and dividing by the fixed nonzero scalar $k!$ are linear operations. For the duality condition 3.112, repeated differentiation gives $\big((x-5)^j\big)^{(k)} = j(j-1)\cdots(j-k+1)(x-5)^{j-k}$ for $k \le j$, and 0 for $k > j$.
(i) $k > j$: the derivative is 0, so $\varphi_k\big((x-5)^j\big) = 0$.
(ii) $k < j$: the factor $(x-5)^{j-k}$ with $j-k \ge 1$ vanishes at 5, so $\varphi_k\big((x-5)^j\big) = 0$.
(iii) $k = j$: the derivative is the constant $k!$, so $\varphi_k\big((x-5)^k\big) = k!/k! = 1$.

Problem 3F.11. Suppose $v_1, \dots, v_n$ is a basis of V and $\varphi_1, \dots, \varphi_n$ is the corresponding dual basis of V'. Suppose $\psi \in V'$. Prove that

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

Solution. Both sides are linear functionals on V agreeing on the basis $v_1, \dots, v_n$, hence are equal by the uniqueness half of the linear map lemma 3.4. Indeed, $\eta := \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n$ lies in V' because V' is a vector space (3.110), and the dual-basis relations $\varphi_k(v_j) = 1$ for $k = j$ and 0 otherwise (3.112) give

$$\eta(v_j) = \sum_{k=1}^{n} \psi(v_k)\varphi_k(v_j) = \psi(v_j)$$

for each j.

Problem 3F.12. Suppose $S, T \in \mathcal{L}(V, W)$.
(a) Prove that $(S+T)' = S' + T'$.
(b) Prove that $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$.
This exercise asks you to verify (a) and (b) in 3.120.

Solution. Both identities come from evaluating at an arbitrary $\varphi \in W'$ and then at an arbitrary $v \in V$, using the definition 3.118 of the dual map, $R'(\varphi) = \varphi \circ R$.
(a)

$$\begin{aligned}\big((S+T)'(\varphi)\big)(v) &= \varphi\big((S+T)v\big) = \varphi(Sv) + \varphi(Tv)\\ &= \big(S'(\varphi)\big)(v) + \big(T'(\varphi)\big)(v) = \big((S'+T')(\varphi)\big)(v),\end{aligned}$$

using 3.5 for $(S+T)v = Sv + Tv$, the additivity of φ, and pointwise addition in V' and in $\mathcal{L}(W', V')$.
(b)

$$\begin{aligned}\big((\lambda T)'(\varphi)\big)(v) &= \varphi(\lambda Tv) = \lambda\,\varphi(Tv)\\ &= \lambda\big(T'(\varphi)\big)(v) = \big((\lambda T')(\varphi)\big)(v),\end{aligned}$$

using 3.5 for $(\lambda T)v = \lambda Tv$, the homogeneity of φ, and pointwise scalar multiplication in V' and in $\mathcal{L}(W', V')$.

Problem 3F.13. Show that the dual map of the identity operator on V is the identity operator on V'.

Solution. For every $\varphi \in V'$, the definition 3.118 of the dual map gives $I'(\varphi) = \varphi \circ I = \varphi$, since

$$(\varphi \circ I)(v) = \varphi(Iv) = \varphi(v) \qquad \text{for every } v \in V.$$

Thus $I' \in \mathcal{L}(V', V')$ fixes every element of V', so it is the identity operator on V'.

Problem 3F.14. Define $T\colon \mathbf{R}^3 \to \mathbf{R}^2$ by

$$T(x, y, z) = (4x + 5y + 6z,\ 7x + 8y + 9z).$$

Suppose φ_1, φ_2 denotes the dual basis of the standard basis of $\mathbf{R}^2$ and ψ_1, ψ_2, ψ_3 denotes the dual

basis of the standard basis of $\mathbf{R}^3$.
(a) Describe the linear functionals $T'(\varphi_1)$ and $T'(\varphi_2)$.
(b) Write $T'(\varphi_1)$ and $T'(\varphi_2)$ as linear combinations of ψ_1, ψ_2, ψ_3.

Solution. (a) $T'(\varphi_1)$ and $T'(\varphi_2)$ are the linear functionals on $\mathbf{R}^3$ given by

$$\big(T'(\varphi_1)\big)(x,y,z) = 4x + 5y + 6z, \qquad \big(T'(\varphi_2)\big)(x,y,z) = 7x + 8y + 9z,$$

that is, the two rows of the formula for T. Indeed $T'(\varphi_j) = \varphi_j \circ T$ by 3.118, and by 3.113 the functionals φ_1, φ_2 select the first and second coordinates of $\mathbf{R}^2$.
(b) By 3.113 again, ψ_1, ψ_2, ψ_3 select the coordinates of $\mathbf{R}^3$, so the functionals $4\psi_1 + 5\psi_2 + 6\psi_3$ and $7\psi_1 + 8\psi_2 + 9\psi_3$ take (x, y, z) to $4x + 5y + 6z$ and to $7x + 8y + 9z$. Comparing with (a),

$$T'(\varphi_1) = 4\psi_1 + 5\psi_2 + 6\psi_3 \qquad \text{and} \qquad T'(\varphi_2) = 7\psi_1 + 8\psi_2 + 9\psi_3.$$

Problem 3F.15. Define $T\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ by

$$(Tp)(x) = x^2 p(x) + p''(x)$$

for each $x \in \mathbb{R}$.
(a) Suppose $\varphi \in \mathcal{P}(\mathbb{R})'$ is defined by $\varphi(p) = p'(4)$. Describe the linear functional $T'(\varphi)$ on $\mathcal{P}(\mathbb{R})$.
(b) Suppose $\varphi \in \mathcal{P}(\mathbb{R})'$ is defined by $\varphi(p) = \int_0^1 p$. Evaluate $(T'(\varphi))(x^3)$.

Solution. (a) $T'(\varphi)$ is the linear functional on $\mathcal{P}(\mathbb{R})$ taking p to $8p(4) + 16\,p'(4) + p'''(4)$. Indeed $(T'(\varphi))(p) = \varphi(Tp) = (Tp)'(4)$ by 3.118, and the product rule gives

$$(Tp)'(x) = 2x\,p(x) + x^2 p'(x) + p'''(x),$$

which at $x = 4$ is the stated value.
(b) $(T'(\varphi))(x^3) = 19/6$. For $p(x) = x^3$ we have $p''(x) = 6x$, so $(Tp)(x) = x^5 + 6x$ and

$$(T'(\varphi))(x^3) = \int_0^1 (x^5 + 6x)\,dx = \left[\frac{x^6}{6} + 3x^2\right]_0^1 = \frac{1}{6} + 3 = \frac{19}{6}.$$

Problem 3F.16. Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$T' = 0 \iff T = 0.$$

Solution. If $T = 0$, then $\big(T'(\varphi)\big)(v) = \varphi(Tv) = \varphi(0) = 0$ for all $\varphi \in W'$ and $v \in V$, so $T' = 0$.
Conversely, suppose $T \neq 0$ and choose $v \in V$ with $Tv \neq 0$. The one-vector list Tv is linearly independent, so since W is finite-dimensional it extends by 2.32 to a basis $w_1 = Tv, w_2, \ldots, w_m$ of W; let $\psi_1, \ldots, \psi_m \in W'$ be its dual basis (3.112). Then

$$\big(T'(\psi_1)\big)(v) = \psi_1(Tv) = \psi_1(w_1) = 1 \neq 0,$$

so $T' \neq 0$. Contrapositively, $T' = 0$ implies $T = 0$.
Method (2) for the second direction: $T' = 0$ gives $\operatorname{null} T' = W'$, and $\operatorname{null} T' = (\operatorname{range} T)^0$ by 3.128(a); since $\operatorname{range} T$ is a subspace of the finite-dimensional space W, 3.127(b) then forces $\operatorname{range} T = \{0\}$, i.e. $T = 0$.

Problem 3F.17. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is invertible if and only if $T' \in \mathcal{L}(W', V')$ is invertible.

Solution. The equivalence is 3.63 applied at both ends of the chain

$$\begin{aligned} T \text{ is invertible} &\iff T \text{ is injective and } T \text{ is surjective} \\ &\iff T' \text{ is surjective and } T' \text{ is injective} \\ &\iff T' \text{ is invertible,} \end{aligned}$$

the middle equivalence being 3.131 (T injective $\iff$ T' surjective) and 3.129 (T surjective $\iff$ T' injective), whose hypothesis that V and W are finite-dimensional holds here.

Problem 3F.18. Suppose V and W are finite-dimensional. Prove that the map that takes $T \in \mathcal{L}(V, W)$ to $T' \in \mathcal{L}(W', V')$ is an isomorphism of $\mathcal{L}(V, W)$ onto $\mathcal{L}(W', V')$.

Solution. The map in question is $\Lambda T = T'$, which lands in $\mathcal{L}(W', V')$ by 3.118; we show it is linear, injective, and of equal-dimensional domain and codomain.
Linearity is 3.120(a) and 3.120(b): for $S, T \in \mathcal{L}(V, W)$ and $\lambda \in \mathbb{F}$,

$$\begin{aligned} \Lambda(S + T) &= S' + T' = \Lambda S + \Lambda T, \\ \Lambda(\lambda T) &= \lambda T' = \lambda\, \Lambda T. \end{aligned}$$

If $\Lambda T = 0$, then $T' = 0$, so $T = 0$ by Exercise 16 of this section (which applies since W is finite-dimensional); hence $\operatorname{null} \Lambda = \{0\}$ and Λ is injective by 3.15. Finally 3.111 gives $\dim V' = \dim V$ and $\dim W' = \dim W$, so 3.72 twice yields

$$\dim \mathcal{L}(W', V') = (\dim W)(\dim V) = \dim \mathcal{L}(V, W).$$

An injective linear map between finite-dimensional spaces of equal dimension is invertible (3.65), so Λ is an isomorphism of $\mathcal{L}(V, W)$ onto $\mathcal{L}(W', V')$.

Problem 3F.19. Suppose $U \subseteq V$. Explain why

$$U^0 = \{\varphi \in V' : U \subseteq \operatorname{null} \varphi\}.$$

Solution. The two conditions on $\varphi \in V'$ are verbatim translations of each other. For fixed $\varphi \in V'$,

$$\begin{aligned} \varphi \in U^0 &\iff \varphi(u) = 0 \text{ for every } u \in U \\ &\iff u \in \operatorname{null} \varphi \text{ for every } u \in U \\ &\iff U \subseteq \operatorname{null} \varphi, \end{aligned}$$

the first step by the definition 3.121 of the annihilator, the second by the definition 3.11 of the null space (each $u \in U$ lies in V, as $U \subseteq V$), and the third by the definition of inclusion. Since both sets are subsets of V', they are equal.

Problem 3F.20. Suppose V is finite-dimensional and U is a subspace of V. Show that

$$U = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}.$$

Solution. Write $X = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}$.
(i) $U \subseteq X$: if $u \in U$ and $\varphi \in U^0$, then $\varphi(u) = 0$ by 3.121.
(ii) $X \subseteq U$, by contraposition. Suppose $v \in V$ and $v \notin U$. The subspace U is finite-dimensional (2.25); let $u_1, \ldots, u_m$ be a basis of it (the empty list if $U = \{0\}$). Then $u_1, \ldots, u_m, v$ is linearly independent, since a relation $a_1 u_1 + \cdots + a_m u_m + bv = 0$ with $b \neq 0$ would put $v \in U$, while $b = 0$ forces $a_1 = \cdots = a_m = 0$. Extend it to a basis

$$u_1, \ldots, u_m,\ v,\ w_1, \ldots, w_k$$

of V (2.32) and let φ be the member of the dual basis (3.112) dual to v, so that

$$\varphi(u_1) = \cdots = \varphi(u_m) = 0, \qquad \varphi(v) = 1.$$

Vanishing on a basis of U, the functional φ vanishes on U, so $\varphi \in U^0$; as $\varphi(v) = 1 \neq 0$, we get $v \notin X$. Hence $U = X$.

Problem 3F.21. Suppose V is finite-dimensional and U and W are subspaces of V.
(a) Prove that $W^0 \subseteq U^0$ if and only if $U \subseteq W$.
(b) Prove that $W^0 = U^0$ if and only if $U = W$.

Solution. (a) Annihilation reverses inclusions. If $U \subseteq W$ and $\varphi \in W^0$, then φ vanishes on W and hence on the subset U, so $\varphi \in U^0$; thus $W^0 \subseteq U^0$.
Conversely, suppose $W^0 \subseteq U^0$ and $u \in U$. Every $\varphi \in W^0$ then lies in U^0, so $\varphi(u) = 0$. Since V is finite-dimensional and W is a subspace of V, Exercise 20 of this section gives

$$W = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in W^0\},$$

so $u \in W$. Hence $U \subseteq W$.
(b) If $U = W$, then trivially $U^0 = W^0$. Conversely, $W^0 = U^0$ gives both $W^0 \subseteq U^0$ and $U^0 \subseteq W^0$, so (a) applied each way gives $U \subseteq W$ and $W \subseteq U$, whence $U = W$.

Problem 3F.22. Suppose V is finite-dimensional and U and W are subspaces of V.
(a) Show that $(U + W)^0 = U^0 \cap W^0$.
(b) Show that $(U \cap W)^0 = U^0 + W^0$.

Solution. (a) Each side consists of the $\varphi \in V'$ vanishing on both U and W. Indeed $U \subseteq U + W$ and $W \subseteq U + W$ (as 0 lies in each), so $\varphi \in (U + W)^0$ forces $\varphi \in U^0 \cap W^0$; conversely if $\varphi \in U^0 \cap W^0$, then for $u \in U$ and $w \in W$,

$$\varphi(u + w) = \varphi(u) + \varphi(w) = 0,$$

so $\varphi \in (U + W)^0$. (Finite-dimensionality is not used here.)
(b) One inclusion is immediate: $\varphi \in U^0$ vanishes on U and hence on $U \cap W$, so $U^0 \subseteq (U \cap W)^0$, and likewise $W^0 \subseteq (U \cap W)^0$; since $(U \cap W)^0$ is a subspace of V' (3.124), it is closed under addition, giving

$$U^0 + W^0 \subseteq (U \cap W)^0.$$

For the dimensions, put $n = \dim V$; then $\dim V' = n$ by 3.111, so 2.43 applies to the subspaces U^0, W^0 of V' and all dimensions below are finite. By 2.43, then (a), then 3.125,

$$\begin{aligned}
\dim(U^0 + W^0) &= \dim U^0 + \dim W^0 - \dim(U^0 \cap W^0) \\
&= \dim U^0 + \dim W^0 - \dim\big((U + W)^0\big) \\
&= (n - \dim U) + (n - \dim W) - \big(n - \dim(U + W)\big) \\
&= n - \dim U - \dim W + \dim(U + W).
\end{aligned}$$

Now 2.43 applied to U and W inside V gives $\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$, and substituting turns the last line into

$$\dim(U^0 + W^0) = n - \dim(U \cap W) = \dim\big((U \cap W)^0\big),$$

the last equality being 3.125 for the subspace $U \cap W$. A subspace of matching finite dimension is the whole space (2.39), so $(U \cap W)^0 = U^0 + W^0$.

Problem 3F.23. Suppose V is finite-dimensional and $\varphi_1, \ldots, \varphi_m \in V'$. Prove that the following three sets are equal to each other.
(a) $\operatorname{span}(\varphi_1, \ldots, \varphi_m)$
(b) $\big((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)\big)^0$
(c) $\{\varphi \in V' : (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m) \subseteq \operatorname{null} \varphi\}$

Solution. Write $W = (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)$, a subspace of V; all three sets equal W^0.
(b) = (c): Exercise 19 of this section says $W^0 = \{\varphi \in V' : W \subseteq \operatorname{null} \varphi\}$, which is (c).
(a) = (b): two ingredients.
(i) $(\operatorname{null} \psi)^0 = \operatorname{span}(\psi)$ for every $\psi \in V'$. If $\psi = 0$, both sides are $\{0\}$ (here $V^0 = \{0\}$ by 3.127(a)). If $\psi \neq 0$, then $\operatorname{range} \psi$ is a nonzero subspace of the one-dimensional space $\mathbf{F}$, so $\dim \operatorname{range} \psi = 1$ and 3.21 then 3.125 give

$$\begin{aligned} \dim \operatorname{null} \psi &= \dim V - 1, \\ \dim (\operatorname{null} \psi)^0 &= 1. \end{aligned}$$

Since ψ vanishes on $\operatorname{null} \psi$, the one-dimensional space $\operatorname{span}(\psi)$ sits inside the one-dimensional space $(\operatorname{null} \psi)^0$, and 2.39 gives equality.
(ii) $(U_1 \cap \cdots \cap U_m)^0 = U_1^0 + \cdots + U_m^0$ for subspaces $U_1, \ldots, U_m$ of V, by induction on m: trivial for $m = 1$, and for $m > 1$ Exercise 22(b) applied to the subspaces $U_1 \cap \cdots \cap U_{m-1}$ and U_m gives

$$(U_1 \cap \cdots \cap U_m)^0 = (U_1 \cap \cdots \cap U_{m-1})^0 + U_m^0,$$

to which the induction hypothesis applies.
Taking $U_k = \operatorname{null} \varphi_k$ in (ii) and then (i) termwise,

$$\begin{aligned} W^0 &= \operatorname{span}(\varphi_1) + \cdots + \operatorname{span}(\varphi_m) \\ &= \operatorname{span}(\varphi_1, \ldots, \varphi_m), \end{aligned}$$

the last step because both sides consist of all $a_1\varphi_1 + \cdots + a_m\varphi_m$ with $a_k \in \mathbf{F}$.

Problem 3F.24. Suppose V is finite-dimensional and $v_1, \ldots, v_m \in V$. Define a linear map $\Gamma : V' \to \mathbf{F}^m$ by $\Gamma(\varphi) = (\varphi(v_1), \ldots, \varphi(v_m))$.
(a) Prove that $v_1, \ldots, v_m$ spans V if and only if Γ is injective.
(b) Prove that $v_1, \ldots, v_m$ is linearly independent if and only if Γ is surjective.

Solution. Both parts rest on $\operatorname{null} \Gamma = U^0$, where $U = \operatorname{span}(v_1, \ldots, v_m)$. Indeed $\Gamma(\varphi) = 0$ says $\varphi(v_k) = 0$ for each k, which forces $\varphi(a_1v_1 + \cdots + a_mv_m) = 0$ by linearity, so $\varphi \in U^0$; conversely $\varphi \in U^0$ kills each $v_k \in U$.
(a) By 3.15, the displayed equality, and 3.127(a),

$$\begin{aligned} \Gamma \text{ is injective} &\iff \operatorname{null} \Gamma = \{0\} \\ &\iff U^0 = \{0\} \iff U = V, \end{aligned}$$

and $U = V$ says exactly that $v_1, \ldots, v_m$ spans V.
(b) By 3.21 applied to $\Gamma \in \mathcal{L}(V', \mathbf{F}^m)$, then 3.111, then $\operatorname{null} \Gamma = U^0$ and 3.125,

$$\begin{aligned} \dim \operatorname{range} \Gamma &= \dim V' - \dim U^0 \\ &= \dim V - \dim U^0 \\ &= \dim V - (\dim V - \dim U) \\ &= \dim U. \end{aligned}$$

Since $\operatorname{range} \Gamma$ is a subspace of the m-dimensional space $\mathbf{F}^m$, 2.39 says Γ is surjective exactly when $\dim \operatorname{range} \Gamma = m$, that is, when $\dim U = m$. Finally $v_1, \ldots, v_m$ is a length-m spanning list of U, so it is linearly independent exactly when $\dim U = m$: independence makes it a basis of U, and $\dim U = m$ makes it a basis by 2.42.

Problem 3F.25. Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m \in V'$. Define a linear map $\Gamma : V \to \mathbf{F}^m$ by $\Gamma(v) = (\varphi_1(v), \dots, \varphi_m(v))$.
(a) Prove that $\varphi_1, \dots, \varphi_m$ spans V' if and only if Γ is injective.
(b) Prove that $\varphi_1, \dots, \varphi_m$ is linearly independent if and only if Γ is surjective.

Solution. Both parts rest on $\operatorname{null}\Gamma = W$, where $W = (\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m)$: indeed $\Gamma(v) = 0$ says $\varphi_k(v) = 0$ for every k, that is, $v \in W$. Exercise 23 of this section supplies

$$W^0 = \operatorname{span}(\varphi_1, \dots, \varphi_m).$$

(a) By 3.15, then $\operatorname{null}\Gamma = W$, then 3.127(b) for the subspace W of V,

$$\begin{aligned} \Gamma \text{ is injective} &\iff \operatorname{null}\Gamma = \{0\} \\ &\iff W = \{0\} \iff W^0 = V', \end{aligned}$$

and by the display $W^0 = V'$ says exactly that $\varphi_1, \dots, \varphi_m$ spans V'.
(b) By 3.21 applied to Γ, then $\operatorname{null}\Gamma = W$ and 3.125,

$$\begin{aligned} \dim \operatorname{range}\Gamma &= \dim V - \dim W = \dim W^0 \\ &= \dim \operatorname{span}(\varphi_1, \dots, \varphi_m), \end{aligned}$$

the last equality again by Exercise 23. Since $\operatorname{range}\Gamma$ is a subspace of the m-dimensional space $\mathbf{F}^m$, 2.39 says Γ is surjective exactly when this dimension equals m. Finally $\varphi_1, \dots, \varphi_m$ is a length-m spanning list of its own span, so its span has dimension m exactly when the list is linearly independent (independence makes it a basis of the span; conversely 2.42 makes it a basis).

Problem 3F.26. Suppose V is finite-dimensional and Ω is a subspace of V'. Prove that

$$\Omega = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Omega\}^0.$$

Solution. Write $W = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Omega\}$; we must show $W^0 = \Omega$.
(i) $\Omega = \{0\}$. Then $W = V$, so $W^0 = V^0 = \{0\} = \Omega$ by 3.127(a).
(ii) $\Omega \neq \{0\}$. Since $\dim V' = \dim V < \infty$ (3.111), the subspace Ω is finite-dimensional (2.25); let $\varphi_1, \dots, \varphi_m$ be a basis of it, $m \geq 1$. Then

$$W = (\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m) :$$

each φ_k lies in Ω, giving $\subseteq$, and if $\varphi_k(v) = 0$ for all k, then every $\varphi = a_1\varphi_1 + \cdots + a_m\varphi_m \in \Omega$ satisfies $\varphi(v) = 0$, giving $\supseteq$. Now Exercise 23 of this section gives

$$W^0 = \operatorname{span}(\varphi_1, \dots, \varphi_m) = \Omega.$$

Problem 3F.27. Suppose $T \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ and $\operatorname{null} T' = \operatorname{span}(\varphi)$, where φ is the linear functional on $\mathcal{P}_5(\mathbf{R})$ defined by $\varphi(p) = p(8)$. Prove that

$$\operatorname{range} T = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}.$$

Solution. The set in question is $\operatorname{null}\varphi$, and both it and $\operatorname{range} T$ turn out to be 5-dimensional subspaces of $V = \mathcal{P}_5(\mathbf{R})$, one inside the other. Here $\dim V = 6$, with basis $1, x, \dots, x^5$.
Since $\varphi(1) = 1$, we have $\varphi \neq 0$, so $\operatorname{range}\varphi$ is a nonzero subspace of $\mathbf{R}$ and hence one-dimensional; also $\dim \operatorname{span}(\varphi) = 1$. By 3.21,

$$\dim \operatorname{null}\varphi = 6 - 1 = 5.$$

By 3.128(a), $(\operatorname{range} T)^0 = \operatorname{null} T' = \operatorname{span}(\varphi)$. In particular φ annihilates $\operatorname{range} T$ (3.121), so $\operatorname{range} T \subseteq$

null φ, and 3.125 applied to the subspace range T of V gives

$$\dim \operatorname{range} T = 6 - \dim \operatorname{span}(\varphi) = 5.$$

By 2.39 inside null φ, therefore range $T = \operatorname{null} \varphi = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}$.

Problem 3F.28. Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m$ is a linearly independent list in V'. Prove that

$$\dim\big((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)\big) = (\dim V) - m.$$

Solution. Put $W = (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)$. Exercise 23 of this section gives $W^0 = \operatorname{span}(\varphi_1, \dots, \varphi_m)$, which has dimension m since the linearly independent list $\varphi_1, \dots, \varphi_m$ is a basis of its span. Hence 3.125 gives

$$\dim W = \dim V - \dim W^0 = (\dim V) - m.$$

Method (2): define $\Gamma \in \mathcal{L}(V, \mathbf{F}^m)$ by $\Gamma(v) = (\varphi_1(v), \dots, \varphi_m(v))$ as in Exercise 25 of this section, so that null $\Gamma = W$ (a vector is killed by Γ exactly when each φ_k kills it). Linear independence of $\varphi_1, \dots, \varphi_m$ makes Γ surjective by Exercise 25(b), so dim range $\Gamma = m$ and 3.21 gives

$$\dim V = \dim W + m.$$

Problem 3F.29. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$.
(a) Prove that if $\varphi \in W'$ and null $T' = \operatorname{span}(\varphi)$, then range $T = \operatorname{null} \varphi$.
(b) Prove that if $\psi \in V'$ and range $T' = \operatorname{span}(\psi)$, then null $T = \operatorname{null} \psi$.

Solution. Both parts are containment plus a dimension count, using throughout that a nonzero $\alpha \in X'$ has range $\alpha = \mathbf{F}$ and hence, by 3.21,

$$\dim \operatorname{null} \alpha = \dim X - 1.$$

(a) By 3.128(a) the hypothesis reads $(\operatorname{range} T)^0 = \operatorname{span}(\varphi)$.
(i) $\varphi = 0$: then $(\operatorname{range} T)^0 = \{0\}$, so range $T = W$ by 3.127(a), while null $\varphi = W$ too.
(ii) $\varphi \neq 0$: then $\varphi \in (\operatorname{range} T)^0$ gives range $T \subseteq \operatorname{null} \varphi$ (3.121), and $\dim \operatorname{span}(\varphi) = 1$, so 3.125 for the subspace range T of W gives

$$\dim \operatorname{range} T = \dim W - 1 = \dim \operatorname{null} \varphi,$$

whence range $T = \operatorname{null} \varphi$ by 2.39.
(b) By 3.130(b) the hypothesis reads $(\operatorname{null} T)^0 = \operatorname{span}(\psi)$.
(i) $\psi = 0$: then $(\operatorname{null} T)^0 = \{0\}$, so null $T = V = \operatorname{null} \psi$ by 3.127(a).
(ii) $\psi \neq 0$: then $\psi \in (\operatorname{null} T)^0$ gives null $T \subseteq \operatorname{null} \psi$, and $\dim \operatorname{span}(\psi) = 1$, so 3.125 for the subspace null T of V gives

$$\dim \operatorname{null} T = \dim V - 1 = \dim \operatorname{null} \psi,$$

whence null $T = \operatorname{null} \psi$ by 2.39.

Problem 3F.30. Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_n$ is a basis of V'. Show that there exists a basis of V whose dual basis is $\varphi_1, \dots, \varphi_n$.

Solution. Take $v_j = \Gamma^{-1}(e_j)$, where $e_1, \dots, e_n$ is the standard basis of $\mathbf{F}^n$ and $\Gamma \in \mathcal{L}(V, \mathbf{F}^n)$ is defined by

$$\Gamma(v) = \big(\varphi_1(v), \dots, \varphi_n(v)\big);$$

here $\dim V = \dim V' = n$ by 3.111.
Γ is invertible. If $v \in \operatorname{null} \Gamma$, then $\varphi_k(v) = 0$ for each k, so every $\varphi = a_1\varphi_1 + \cdots + a_n\varphi_n \in V'$ satisfies

$\varphi(v) = 0$ (the list spans V'); were $v \neq 0$, extending v to a basis of V (2.32) and applying the linear map lemma (3.4) would produce $\varphi \in V'$ with $\varphi(v) = 1$, a contradiction. So $\operatorname{null}\Gamma = \{0\}$, and Γ is injective by 3.15, hence invertible by 3.65 and 3.63 since $\dim V = n = \dim \mathbf{F}^n$.
Then $v_1, \ldots, v_n$ is linearly independent: applying Γ to $a_1v_1 + \cdots + a_nv_n = 0$ gives $a_1e_1 + \cdots + a_ne_n = 0$, so all $a_j = 0$. Being independent of length $n = \dim V$, it is a basis of V (2.38). Finally $\Gamma(v_j) = e_j$ says $\varphi_k(v_j)$ is 1 for $k = j$ and 0 otherwise, which by 3.112 is exactly the statement that $\varphi_1, \ldots, \varphi_n$ is the dual basis of $v_1, \ldots, v_n$.

Problem 3F.31. Suppose U is a subspace of V. Let $i\colon U \to V$ be the inclusion map defined by $i(u) = u$. Thus $i' \in \mathcal{L}(V', U')$.
(a) Show that $\operatorname{null} i' = U^0$.
(b) Prove that if V is finite-dimensional, then $\operatorname{range} i' = U'$.
(c) Prove that if V is finite-dimensional, then $\widetilde{i'}$ is an isomorphism from V'/U^0 onto U'.
[The isomorphism in (c) is natural in that it does not depend on a choice of basis in either vector space.]

Solution. Everything follows from $i'(\varphi) = \varphi \circ i = \varphi|_U$, the restriction of φ to U, by the definition 3.118 of the dual map.
(a) For $\varphi \in V'$, $\varphi \in \operatorname{null} i'$ says $\varphi|_U = 0$, that is, $\varphi(u) = 0$ for every $u \in U$, which is $\varphi \in U^0$ by 3.121. (No finite-dimensionality needed.)
(b) One inclusion holds since $i' \in \mathcal{L}(V', U')$. Conversely, let $\varphi \in U'$. The subspace U is finite-dimensional (2.25); take a basis $u_1, \ldots, u_m$ of U and extend to a basis $u_1, \ldots, u_n$ of V (2.32). The linear map lemma (3.4) gives $\psi \in V'$ with

$$\psi(u_j) = \begin{cases} \varphi(u_j) & \text{if } 1 \leq j \leq m, \\ 0 & \text{if } m < j \leq n. \end{cases}$$

Then $\psi|_U$ and φ agree on a basis of U, hence on U, so $i'(\psi) = \varphi$.
(c) In the notation of 3.106, $\widetilde{i'}$ maps $V'/(\operatorname{null} i') = V'/U^0$ to U' by $\varphi + U^0 \mapsto i'(\varphi)$, the domain being identified via (a). It is injective by 3.107(b), and surjective because $\operatorname{range} \widetilde{i'} = \operatorname{range} i' = U'$ by 3.107(c) and (b). Hence it is invertible (3.63), so an isomorphism from V'/U^0 onto U'.

Problem 3F.32. The double dual space of V, denoted by V'', is defined to be the dual space of V'. In other words, $V'' = (V')'$. Define $\Lambda\colon V \to V''$ by

$$(\Lambda v)(\varphi) = \varphi(v)$$

for each $v \in V$ and each $\varphi \in V'$.
(a) Show that Λ is a linear map from V to V''.
(b) Show that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$, where $T'' = (T')'$.
(c) Show that if V is finite-dimensional, then Λ is an isomorphism from V onto V''.
[Suppose V is finite-dimensional. Then V and V' are isomorphic, but finding an isomorphism from V onto V' generally requires choosing a basis of V. In contrast, the isomorphism Λ from V onto V'' does not require a choice of basis and thus is considered more natural.]

Solution. (a) Each Λv is linear on V', because for $\varphi, \psi \in V'$ and $\lambda \in \mathbf{F}$ the definitions of addition and scalar multiplication in $V' = \mathcal{L}(V, \mathbf{F})$ give

$$(\Lambda v)(\varphi + \psi) = (\varphi + \psi)(v) = \varphi(v) + \psi(v),$$
$$(\Lambda v)(\lambda\varphi) = (\lambda\varphi)(v) = \lambda\, \varphi(v);$$

these are $(\Lambda v)(\varphi) + (\Lambda v)(\psi)$ and $\lambda(\Lambda v)(\varphi)$, so $\Lambda v \in V''$. And Λ itself is linear, since for all $u, v \in V$,

$\lambda \in \mathbf{F}$ and every $\varphi \in V'$,

$$\begin{aligned}\big(\Lambda(u+v)\big)(\varphi) &= \varphi(u) + \varphi(v),\\ \big(\Lambda(\lambda v)\big)(\varphi) &= \lambda\,\varphi(v),\end{aligned}$$

which are the values at φ of $\Lambda u + \Lambda v$ and of $\lambda\Lambda v$. Thus $\Lambda \in \mathcal{L}(V, V'')$.
(b) Both $T'' \circ \Lambda$ and $\Lambda \circ T$ map V to V'', so fix $v \in V$ and $\varphi \in V'$ and compare values at φ. Using the definition 3.118 of the dual map twice,

$$\begin{aligned}\big((T'' \circ \Lambda)(v)\big)(\varphi) &= \big(T''(\Lambda v)\big)(\varphi)\\ &= \big((\Lambda v) \circ T'\big)(\varphi)\\ &= (\Lambda v)\big(T'(\varphi)\big)\\ &= (\Lambda v)(\varphi \circ T)\\ &= (\varphi \circ T)(v)\\ &= \varphi(Tv)\\ &= \big(\Lambda(Tv)\big)(\varphi)\\ &= \big((\Lambda \circ T)(v)\big)(\varphi),\end{aligned}$$

where the second equality uses $T''(\alpha) = \alpha \circ T'$, the fourth uses $T'(\varphi) = \varphi \circ T$, and the fifth and seventh use the definition of Λ. As φ and v were arbitrary, $T'' \circ \Lambda = \Lambda \circ T$.
(c) Applying 3.111 twice gives $\dim V'' = \dim V' = \dim V < \infty$. If $v \in \operatorname{null} \Lambda$, then $\varphi(v) = 0$ for every $\varphi \in V'$; were $v \neq 0$, extending v to a basis of V (2.32) and applying the linear map lemma (3.4) would give $\varphi \in V'$ with $\varphi(v) = 1$, a contradiction. So $\operatorname{null} \Lambda = \{0\}$ and Λ is injective by 3.15, hence surjective by 3.65 and invertible by 3.63. Thus Λ is an isomorphism from V onto V''.

Problem 3F.33. Suppose U is a subspace of V. Let $\pi\colon V \to V/U$ be the usual quotient map. Thus $\pi' \in \mathcal{L}\big((V/U)', V'\big)$.
(a) Show that π' is injective.
(b) Show that $\operatorname{range} \pi' = U^0$.
(c) Conclude that π' is an isomorphism from $(V/U)'$ onto U^0.
[The isomorphism in (c) is natural in that it does not depend on a choice of basis in either vector space. In fact, there is no assumption here that any of these vector spaces are finite-dimensional.]

Solution. All three parts run off the formula $\big(\pi'(\varphi)\big)(v) = \varphi(v + U)$, from $\pi(v) = v + U$ and the definition 3.118 of the dual map. No finite-dimensionality is used.
(a) If $\varphi \in \operatorname{null} \pi'$, then $\varphi(v + U) = 0$ for every $v \in V$; since π is surjective, this says φ vanishes on V/U, so $\varphi = 0$. Hence $\operatorname{null} \pi' = \{0\}$ and 3.15 gives injectivity.
(b) For $\varphi \in (V/U)'$ and $u \in U$ we have $u + U = 0 + U$ by 3.101, so

$$\big(\pi'(\varphi)\big)(u) = \varphi(0 + U) = 0$$

($0 + U$ is the additive identity of V/U and φ is linear, 3.10); thus $\pi'(\varphi) \in U^0$ by 3.121.
Conversely, let $\psi \in U^0$ and define $\varphi\colon V/U \to \mathbf{F}$ by $\varphi(v + U) = \psi(v)$. This is well defined, since $v + U = w + U$ gives $v - w \in U$ (3.101) and hence $\psi(v) - \psi(w) = \psi(v - w) = 0$; and it is linear, by the definitions of the operations on V/U together with linearity of ψ (Check!). Then $\big(\pi'(\varphi)\big)(v) = \psi(v)$ for all v, so $\psi \in \operatorname{range} \pi'$. Hence $\operatorname{range} \pi' = U^0$.
(c) By 3.124, U^0 is a subspace of V', and by (b) we may regard π' as a map into U^0; it is injective by (a) and surjective onto U^0 by (b), hence invertible (3.63). So π' is an isomorphism from $(V/U)'$ onto U^0.

4 Polynomials

Contents

4.1 Exercises

Problem 4.1. Suppose $w, z \in \mathbb{C}$. Verify the following equalities and inequalities.
(a) $z + \bar{z} = 2 \operatorname{Re} z$
(b) $z - \bar{z} = 2(\operatorname{Im} z)i$
(c) $z\bar{z} = |z|^2$
(d) $\overline{w+z} = \bar{w} + \bar{z}$ and $\overline{wz} = \bar{w}\,\bar{z}$
(e) $\bar{\bar{z}} = z$
(f) $|\operatorname{Re} z| \le |z|$ and $|\operatorname{Im} z| \le |z|$
(g) $|\bar{z}| = |z|$
(h) $|wz| = |w|\,|z|$
The results above are the parts of 4.4 that were left to the reader.

Solution. Write $z = a + bi$ and $w = c + di$ with $a, b, c, d \in \mathbb{R}$, so that by 4.2

$$\begin{aligned} \bar{z} &= a - bi, & |z| &= \sqrt{a^2+b^2}, \\ \bar{w} &= c - di, & |w| &= \sqrt{c^2+d^2}. \end{aligned}$$

(a) $z + \bar{z} = (a + bi) + (a - bi) = 2a = 2 \operatorname{Re} z$.
(b) $z - \bar{z} = (a + bi) - (a - bi) = 2bi = 2(\operatorname{Im} z)i$.
(c) Using $i^2 = -1$,

$$z\bar{z} = (a+bi)(a-bi) = a^2 + b^2 = |z|^2.$$

(d) Since $w + z = (c + a) + (d + b)i$ with real parts and imaginary parts as shown,

$$\overline{w+z} = (c+a) - (d+b)i = (c-di) + (a-bi) = \bar{w} + \bar{z}.$$

And $wz = (ca - db) + (cb + da)i$ with $ca - db, cb + da \in \mathbb{R}$, so

$$\overline{wz} = (ca-db) - (cb+da)i = (c-di)(a-bi) = \bar{w}\,\bar{z}.$$

(e) $\bar{z} = a + (-b)i$ with $a, -b \in \mathbb{R}$, so $\bar{\bar{z}} = a - (-b)i = a + bi = z$.
(f) From $b^2 \ge 0$ and $a^2 \ge 0$, with the square root increasing on $[0, \infty)$,

$$\begin{aligned} |\operatorname{Re} z| &= \sqrt{a^2} \le \sqrt{a^2+b^2} = |z|, \\ |\operatorname{Im} z| &= \sqrt{b^2} \le \sqrt{a^2+b^2} = |z|. \end{aligned}$$

(g) By (e), $\operatorname{Re} \bar{z} = a$ and $\operatorname{Im} \bar{z} = -b$, so $|\bar{z}| = \sqrt{a^2+b^2} = |z|$.
(h) By (c) three times, (d) once, and commutativity,

$$\begin{aligned} |wz|^2 &= (wz)\overline{(wz)} = (w\bar{w})(z\bar{z}) \\ &= |w|^2|z|^2 = \big(|w|\,|z|\big)^2, \end{aligned}$$

and both $|wz|$ and $|w|\,|z|$ are nonnegative, so $|wz| = |w|\,|z|$.

Problem 4.2. Prove that if $w, z \in \mathbb{C}$, then $\big|\,|w| - |z|\,\big| \le |w - z|$.
The inequality above is called the reverse triangle inequality.

Solution. The triangle inequality (4.4) applied to $w = (w - z) + z$ gives

$$|w| \le |w-z| + |z|, \qquad \text{so} \qquad |w| - |z| \le |w-z|.$$

Interchanging w and z gives $|z| - |w| \le |z - w| = |w - z|$, the last equality because multiplicativity of absolute value (4.4, Exercise 4.1(h)) gives $|-u| = |-1|\,|u| = |u|$, as $|-1| = 1$. Since $\big|\,|w| - |z|\,\big|$ is one of the two left sides, it is at most $|w - z|$.

Problem 4.3. Suppose V is a complex vector space and $\varphi \in V'$. Define $\sigma\colon V \to \mathbb{R}$ by $\sigma(v) = \operatorname{Re}\varphi(v)$ for each $v \in V$. Show that

$$\varphi(v) = \sigma(v) - i\sigma(iv)$$

for all $v \in V$.

Solution. Fix $v \in V$ and write $\varphi(v) = \alpha + \beta i$ with $\alpha, \beta \in \mathbb{R}$, so $\sigma(v) = \alpha$. Complex homogeneity of φ at the scalar i gives

$$\varphi(iv) = i\varphi(v) = i(\alpha + \beta i) = -\beta + \alpha i,$$

whose real part is $-\beta$ since $-\beta, \alpha \in \mathbb{R}$; thus $\sigma(iv) = -\beta$ and

$$\sigma(v) - i\sigma(iv) = \alpha - i(-\beta) = \alpha + \beta i = \varphi(v).$$

Problem 4.4. Suppose m is a positive integer. Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbb{F}) : \deg p = m\}$$

a subspace of $\mathcal{P}(\mathbb{F})$?

Solution. No: take $p(z) = z^m$ and $q(z) = 1 - z^m$. Both lie in the set U in question, since each has z^m-coefficient $\pm 1 \neq 0$ and no higher power, so $\deg p = \deg q = m$. But

$$(p+q)(z) = z^m + (1 - z^m) = 1,$$

the nonzero constant polynomial, of degree $0 \neq m$ since m is positive. So $p+q \notin U$ and U is not closed under addition, hence not a subspace of $\mathcal{P}(\mathbb{F})$.

Problem 4.5. Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbb{F}) : \deg p \text{ is even}\}$$

a subspace of $\mathcal{P}(\mathbb{F})$?

Solution. No: take $p(z) = z^2 + z$ and $q(z) = -z^2$. Both have degree 2, so both lie in the set U in question, but

$$(p+q)(z) = (z^2 + z) + (-z^2) = z$$

is nonzero of odd degree 1. So $p + q \notin U$, and as in Exercise 4.4 closure under addition fails, hence U is not a subspace of $\mathcal{P}(\mathbb{F})$.

Problem 4.6. Suppose that m and n are positive integers with $m \leq n$, and suppose $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbb{F})$ with $\deg p = n$ such that $0 = p(\lambda_1) = \cdots = p(\lambda_m)$ and such that p has no other zeros.

Solution. Take

$$p(z) = (z - \lambda_1)^{n-m+1}(z - \lambda_2)\cdots(z - \lambda_m),$$

which reads $p(z) = (z - \lambda_1)^n$ when $m = 1$; the exponent $n - m + 1$ is a positive integer because $m \leq n$.
Degree: the n factors are each monic of degree 1, and a product of monic polynomials is monic with degree the sum of the degrees (multiplying out produces the top power with coefficient 1 and nothing higher), so $\deg p = (n - m + 1) + (m - 1) = n$; in particular $p \neq 0$.
Zeros: $p(\lambda_k) = 0$ for each k, since the factor $\lambda_k - \lambda_k$ occurs. Conversely, if $p(\lambda) = 0$, then because $\mathbb{F}$ has no zero divisors some factor $\lambda - \lambda_k$ vanishes, so $\lambda \in \{\lambda_1, \ldots, \lambda_m\}$. Thus the zero set of p is exactly $\{\lambda_1, \ldots, \lambda_m\}$.

Problem 4.7. Suppose that m is a nonnegative integer, $z_1, \ldots, z_{m+1}$ are distinct elements of $\mathbb{F}$, and $w_1, \ldots, w_{m+1} \in \mathbb{F}$. Prove that there exists a unique polynomial $p \in \mathcal{P}_m(\mathbb{F})$ such that

$$p(z_k) = w_k$$

for each $k = 1, \ldots, m+1$.
This result can be proved without using linear algebra. However, try to find the clearer, shorter proof that uses some linear algebra.

Solution. The evaluation map $T\colon \mathcal{P}_m(\mathbb{F}) \to \mathbb{F}^{m+1}$,

$$Tp = \big(p(z_1), \ldots, p(z_{m+1})\big),$$

is invertible; existence and uniqueness are then its surjectivity and injectivity at $(w_1, \ldots, w_{m+1})$. It is linear because addition and scalar multiplication of functions are pointwise.
T is injective. If $Tp = 0$, then p vanishes at the $m+1$ distinct points $z_1, \ldots, z_{m+1}$. Were $p \neq 0$, then either $\deg p = 0$, so p is a nonzero constant with no zeros at all, or $\deg p = k$ with $1 \leq k \leq m$, and 4.8 allows p at most $k \leq m$ zeros; both contradict having $m+1$ zeros. So $\operatorname{null} T = \{0\}$ and 3.15 applies.
Since $1, z, \ldots, z^m$ is a basis of $\mathcal{P}_m(\mathbb{F})$ (2.36),

$$\dim \mathcal{P}_m(\mathbb{F}) = m + 1 = \dim \mathbb{F}^{m+1},$$

so the injective map T between finite-dimensional spaces of equal dimension is invertible by 3.65. Hence there is exactly one $p \in \mathcal{P}_m(\mathbb{F})$ with $Tp = (w_1, \ldots, w_{m+1})$, that is, with $p(z_k) = w_k$ for each k.

Problem 4.8. Suppose $p \in \mathcal{P}(\mathbb{C})$ has degree m. Prove that p has m distinct zeros if and only if p and its derivative p' have no zeros in common.

Solution. The equivalence is read off the factorization $p(z) = c(z - \lambda_1) \cdots (z - \lambda_m)$ with $c \neq 0$ supplied by 4.13, whose zero set is $\{\lambda_1, \ldots, \lambda_m\}$; so p has m distinct zeros exactly when $\lambda_1, \ldots, \lambda_m$ are distinct.
Throughout, the product rule $(fg)' = f'g + fg'$ holds on $\mathcal{P}(\mathbb{C})$: both sides are bilinear in (f, g), so it suffices to check monomials, where it reads $(j+k)z^{j+k-1} = jz^{j-1}z^k + z^j kz^{k-1}$.
If $m = 0$, then p is a nonzero constant with no zeros at all, so p has $m = 0$ distinct zeros and shares no zero with p'; both sides hold. Assume $m \geq 1$.
(i) $\lambda_1, \ldots, \lambda_m$ distinct. The product rule applied repeatedly gives

$$p'(z) = c \sum_{k=1}^{m} \prod_{j \neq k} (z - \lambda_j),$$

and at $z = \lambda_k$ every summand but the kth contains the factor $\lambda_k - \lambda_k = 0$, leaving

$$p'(\lambda_k) = c \prod_{j \neq k} (\lambda_k - \lambda_j) \neq 0$$

since $c \neq 0$ and the λ's are distinct. So no zero of p is a zero of p'.
(ii) Two of the λ's coincide, say $\lambda_1 = \lambda_2 =: \lambda$. With $g(z) = c(z - \lambda_3) \cdots (z - \lambda_m)$ we get $p(z) = (z - \lambda)^2 g(z)$, so the product rule gives

$$p'(z) = 2(z - \lambda)g(z) + (z - \lambda)^2 g'(z),$$

whence $p'(\lambda) = 0 = p(\lambda)$ and λ is a common zero. Contrapositively, no common zero forces m distinct zeros.

Problem 4.9. Prove that every polynomial of odd degree with real coefficients has a real zero.

Solution. The zero is λ_1, the first linear factor in the real factorization. Suppose $p \in \mathcal{P}(\mathbb{R})$ has odd degree n, so $n \geq 1$ and p is nonconstant; by 4.16 there are reals $c, \lambda_1, \dots, \lambda_m, b_1, \dots, b_M, c_1, \dots, c_M$ with $b_k^2 < 4c_k$ for each k and

$$\begin{aligned} p(x) = c(x - \lambda_1) \cdots (x - \lambda_m) \\ \times (x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M). \end{aligned}$$

Here $c \neq 0$, since otherwise the right side would be the zero polynomial. Comparing degrees,

$$n = m + 2M,$$

so m is odd because n is; in particular $m \geq 1$, the factor $x - \lambda_1$ is present, and $p(\lambda_1) = 0$.

Problem 4.10. For $p \in \mathcal{P}(\mathbb{R})$, define $Tp\colon \mathbb{R} \to \mathbb{R}$ by

$$(Tp)(x) = \begin{cases} \dfrac{p(x) - p(3)}{x - 3} & \text{if } x \neq 3, \\ p'(3) & \text{if } x = 3 \end{cases}$$

for each $x \in \mathbb{R}$. Show that $Tp \in \mathcal{P}(\mathbb{R})$ for every polynomial $p \in \mathcal{P}(\mathbb{R})$ and also show that $T\colon \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$ is a linear map.

Solution. Tp is the quotient q in the factorization $p(x) - p(3) = (x - 3)q(x)$.
(i) If p is constant then $(Tp)(x) = 0$ for $x \neq 3$ and $(Tp)(3) = p'(3) = 0$, so $Tp = 0 \in \mathcal{P}(\mathbb{R})$.
(ii) If $\deg p = m \geq 1$, then $x \mapsto p(x) - p(3)$ has degree m and vanishes at 3, so 4.6 supplies $q \in \mathcal{P}(\mathbb{R})$ with $\deg q = m - 1$ and

$$p(x) - p(3) = (x - 3)q(x) \qquad \text{for all } x \in \mathbb{R}.$$

Dividing gives $(Tp)(x) = q(x)$ for $x \neq 3$; differentiating gives $p'(x) = q(x) + (x - 3)q'(x)$, so $x = 3$ yields $(Tp)(3) = p'(3) = q(3)$. Hence $Tp = q \in \mathcal{P}(\mathbb{R})$.
For linearity, fix $p_1, p_2 \in \mathcal{P}(\mathbb{R})$ and $\alpha \in \mathbb{R}$ and compare values at each x. For $x \neq 3$,

$$\begin{aligned} \big(T(\alpha p_1 + p_2)\big)(x) &= \frac{\alpha p_1(x) + p_2(x) - \alpha p_1(3) - p_2(3)}{x - 3} \\ &= \alpha (Tp_1)(x) + (Tp_2)(x), \end{aligned}$$

while at $x = 3$, since differentiation is linear,

$$\big(T(\alpha p_1 + p_2)\big)(3) = \alpha p_1'(3) + p_2'(3) = \alpha (Tp_1)(3) + (Tp_2)(3).$$

Thus $T(\alpha p_1 + p_2) = \alpha\, Tp_1 + Tp_2$, so T is linear.

Problem 4.11. Suppose $p \in \mathcal{P}(\mathbb{C})$. Define $q\colon \mathbb{C} \to \mathbb{C}$ by

$$q(z) = p(z)\, \overline{p(\bar{z})}.$$

Prove that q is a polynomial with real coefficients.

Solution. Writing $p(z) = \sum_{k=0}^{m} a_k z^k$ and $\bar{p}(z) = \sum_{k=0}^{m} \overline{a_k} z^k$, we have $q = p\,\bar{p}$, a product of polynomials and hence a polynomial.

Indeed, conjugation is additive and multiplicative and $\overline{\overline{z}} = z$ (4.4), so

$$\overline{p(\overline{z})} = \sum_{k=0}^{m} \overline{a_k}\left(\overline{\overline{z}}\right)^k = \sum_{k=0}^{m} \overline{a_k} z^k = \bar{p}(z).$$

Expanding the product and collecting powers of z,

$$q(z) = \sum_{n=0}^{2m} c_n z^n, \qquad c_n = \sum_{\substack{j+k=n \\ 0 \le j,\, k \le m}} a_j \overline{a_k},$$

and by the uniqueness of coefficients (a consequence of 4.8) these are the coefficients of q. The swap $(j,k) \mapsto (k,j)$ carries the index set to itself, so conjugating and relabeling gives

$$\overline{c_n} = \sum_{j+k=n} \overline{a_j} a_k = \sum_{j+k=n} a_j \overline{a_k} = c_n,$$

whence $c_n \in \mathbb{R}$ (4.4) for every n. Thus q is a polynomial with real coefficients.

Problem 4.12. Suppose m is a nonnegative integer and $p \in \mathcal{P}_m(\mathbb{C})$ is such that there are distinct real numbers $x_0, x_1, \ldots, x_m$ with $p(x_k) \in \mathbb{R}$ for each $k = 0, 1, \ldots, m$. Prove that all coefficients of p are real.

Solution. The polynomial of imaginary parts vanishes identically. Write $p(z) = \sum_{k=0}^{m} a_k z^k$ with $a_k \in \mathbb{C}$ and set

$$u(x) = \sum_{k=0}^{m} (\operatorname{Im} a_k) x^k \in \mathcal{P}_m(\mathbb{R}).$$

For $x \in \mathbb{R}$ each x^k is real, and Im is additive with $\operatorname{Im}(at) = t \operatorname{Im} a$ for $t \in \mathbb{R}$, so

$$\operatorname{Im} p(x) = \sum_{k=0}^{m} \operatorname{Im}\left(a_k x^k\right) = u(x).$$

By hypothesis $u(x_0) = u(x_1) = \cdots = u(x_m) = 0$, so u has $m+1$ distinct zeros. Were $u \neq 0$, its degree d would satisfy $0 \le d \le m$, and u would have at most $d \le m$ zeros (none if $d = 0$; by 4.8 if $d \ge 1$) – a contradiction. Hence $u = 0$, so by uniqueness of coefficients $\operatorname{Im} a_k = 0$ for every k; all coefficients of p are real.
Method (2): put $w_k = p(x_k) \in \mathbb{R}$. Exercise 4.7 over $\mathbb{R}$ gives $\tilde{p} \in \mathcal{P}_m(\mathbb{R})$ with $\tilde{p}(x_k) = w_k$ for each k; viewing $\tilde{p} \in \mathcal{P}_m(\mathbb{C})$, the uniqueness assertion of Exercise 4.7 over $\mathbb{C}$ forces $p = \tilde{p}$.

Problem 4.13. Suppose $p \in \mathcal{P}(\mathbb{F})$ with $p \neq 0$. Let $U = \{pq : q \in \mathcal{P}(\mathbb{F})\}$.
(a) Show that $\dim \mathcal{P}(\mathbb{F})/U = \deg p$.
(b) Find a basis of $\mathcal{P}(\mathbb{F})/U$.

Solution. With $m = \deg p$, the cosets

$$1 + U,\ z + U,\ z^2 + U,\ \ldots,\ z^{m-1} + U$$

form a basis of $\mathcal{P}(\mathbb{F})/U$, whence $\dim \mathcal{P}(\mathbb{F})/U = m = \deg p$; this answers (b) and (a). (For $m = 0$ the list is empty and $U = \mathcal{P}(\mathbb{F})$.) Here U is indeed a subspace, since $0 = p \cdot 0$, $pq_1 + pq_2 = p(q_1 + q_2)$ and $\lambda(pq) = p(\lambda q)$.
Spanning: given $r \in \mathcal{P}(\mathbb{F})$, the division algorithm 4.9 (legitimate as $p \neq 0$) yields $q, s \in \mathcal{P}(\mathbb{F})$ with $r = pq + s$ and $\deg s < m$, so $s = a_0 + a_1 z + \cdots + a_{m-1} z^{m-1}$; as $pq \in U$ and the quotient map is linear (3.104),

$$r + U = s + U = a_0(1 + U) + a_1(z + U) + \cdots + a_{m-1}\left(z^{m-1} + U\right).$$

Independence: if that combination equals $0+U$, then $s = a_0 + \cdots + a_{m-1}z^{m-1}$ lies in U by 3.101, say $s = pq$. Were $q \neq 0$ we would get $\deg s = m + \deg q \geq m$, contradicting $\deg s \leq m-1$; so $q = 0$, hence $s = 0$, and uniqueness of coefficients (4.8) forces $a_0 = \cdots = a_{m-1} = 0$.

Problem 4.14. Suppose $p, q \in \mathcal{P}(\mathbb{C})$ are nonconstant polynomials with no zeros in common. Let $m = \deg p$ and $n = \deg q$. Use linear algebra as outlined below in (a)–(c) to prove that there exist $r \in \mathcal{P}_{n-1}(\mathbb{C})$ and $s \in \mathcal{P}_{m-1}(\mathbb{C})$ such that

$$rp + sq = 1.$$

(a) Define $T\colon \mathcal{P}_{n-1}(\mathbb{C}) \times \mathcal{P}_{m-1}(\mathbb{C}) \to \mathcal{P}_{m+n-1}(\mathbb{C})$ by

$$T(r,s) = rp + sq.$$

Show that the linear map T is injective.
(b) Show that the linear map T in (a) is surjective.
(c) Use (b) to conclude that there exist $r \in \mathcal{P}_{n-1}(\mathbb{C})$ and $s \in \mathcal{P}_{m-1}(\mathbb{C})$ such that $rp + sq = 1$.

Solution. Here $m, n \geq 1$, and T does land in $\mathcal{P}_{m+n-1}(\mathbb{C})$ because $\deg(rp) \leq (n-1)+m$ and $\deg(sq) \leq (m-1)+n$. Throughout we use $\deg(uv) = \deg u + \deg v$, which makes $\mathcal{P}(\mathbb{C})$ free of zero divisors and so permits cancellation of a nonzero factor.
(a) Suppose $T(r,s) = 0$, that is, $rp = -sq$. Since q is nonconstant, 4.13 gives $c \neq 0$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ (repetitions allowed) with

$$q(z) = c(z-\lambda_1)\cdots(z-\lambda_n),$$

and each λ_k, being a zero of q, satisfies $p(\lambda_k) \neq 0$. Assume $r \neq 0$; we show by induction on j that for $j = 0, 1, \ldots, n$ there is $r_j \neq 0$ with

$$r(z) = (z-\lambda_1)\cdots(z-\lambda_j)\, r_j(z).$$

Take $r_0 = r$. Given r_j with $j < n$, substituting this factorization of r into $rp = -sq$ and cancelling the nonzero factor $(z-\lambda_1)\cdots(z-\lambda_j)$ leaves

$$r_j(z)p(z) = -c\, s(z)\,(z-\lambda_{j+1})\cdots(z-\lambda_n);$$

evaluating at $z = \lambda_{j+1}$ kills the right side, and $p(\lambda_{j+1}) \neq 0$ then forces $r_j(\lambda_{j+1}) = 0$, so 4.6 yields r_{j+1} of degree $\deg r_j - 1$ (hence nonzero) with $r_j(z) = (z-\lambda_{j+1})r_{j+1}(z)$. At $j = n$ this gives $\deg r = n + \deg r_n \geq n$, contradicting $\deg r \leq n-1$. So $r = 0$, whence $sq = 0$ with $q \neq 0$ gives $s = 0$. Thus $\operatorname{null} T = \{(0,0)\}$ and T is injective (3.15).
(b) Since $\dim \mathcal{P}_k(\mathbb{C}) = k+1$ and dimension adds over products (3.92),

$$\dim\bigl(\mathcal{P}_{n-1}(\mathbb{C}) \times \mathcal{P}_{m-1}(\mathbb{C})\bigr) = n + m = \dim \mathcal{P}_{m+n-1}(\mathbb{C}),$$

so the injective T between spaces of equal finite dimension is surjective by 3.65.
(c) The constant polynomial 1 lies in $\mathcal{P}_{m+n-1}(\mathbb{C})$, so by (b) there is a pair $(r,s) \in \mathcal{P}_{n-1}(\mathbb{C}) \times \mathcal{P}_{m-1}(\mathbb{C})$ with $rp + sq = T(r,s) = 1$.

5 Eigenvalues and Eigenvectors

Contents

5.1 Exercises 5A

Problem 5A.1. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V.
(a) Prove that if $U \subseteq \operatorname{null} T$, then U is invariant under T.
(b) Prove that if $\operatorname{range} T \subseteq U$, then U is invariant under T.

Solution. In both parts take $u \in U$; invariance (5.2) asks that $Tu \in U$.
(a) If $U \subseteq \operatorname{null} T$, then $Tu = 0 \in U$, since the subspace U contains 0.
(b) If $\operatorname{range} T \subseteq U$, then $Tu \in \operatorname{range} T \subseteq U$.

Problem 5A.2. Suppose that $T \in \mathcal{L}(V)$ and $V_1, \ldots, V_m$ are subspaces of V invariant under T. Prove that $V_1 + \cdots + V_m$ is invariant under T.

Solution. T sends a decomposition of v to a decomposition of Tv. Let $v \in V_1 + \cdots + V_m$ (a subspace, by 1.40) and write $v = v_1 + \cdots + v_m$ with $v_k \in V_k$. By additivity of T,

$$Tv = Tv_1 + \cdots + Tv_m,$$

and $Tv_k \in V_k$ for each k since V_k is invariant. Hence $Tv \in V_1 + \cdots + V_m$.

Problem 5A.3. Suppose $T \in \mathcal{L}(V)$. Prove that the intersection of every collection of subspaces of V invariant under T is invariant under T.

Solution. Membership in an intersection is checked one index at a time. Let $\{U_a\}_{a \in A}$ be such a collection and $U = \bigcap_{a \in A} U_a$. Then U is a subspace by 1.34: it contains 0, and $u, w \in U$, $\lambda \in \mathbf{F}$ give $u + w, \lambda u \in U_a$ for every a, hence in U. Now for $u \in U$ and any $a \in A$ we have $u \in U_a$, so invariance of U_a gives $Tu \in U_a$; as a was arbitrary,

$$Tu \in \bigcap_{a \in A} U_a = U.$$

Problem 5A.4. Prove or give a counterexample: If V is finite-dimensional and U is a subspace of V that is invariant under every operator on V, then $U = \{0\}$ or $U = V$.

Solution. True. Suppose $U \neq \{0\}$ and pick $u \in U$ with $u \neq 0$; the one-vector list u is then linearly independent, so 2.32 extends it to a basis $u, e_2, \ldots, e_n$ of V. Given any $v \in V$, the linear map lemma 3.4 supplies $T \in \mathcal{L}(V)$ with

$$Tu = v, \qquad Te_2 = \cdots = Te_n = 0.$$

By hypothesis U is invariant under this T, and $u \in U$, so $v = Tu \in U$. Hence $V \subseteq U$, that is, $U = V$.

Problem 5A.5. Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.

Solution. T has no eigenvalues. Indeed, $T(x, y) = \lambda(x, y)$ says

$$-3y = \lambda x, \qquad x = \lambda y,$$

and substituting the second into the first gives $(\lambda^2 + 3)y = 0$. If $y = 0$ then $x = \lambda y = 0$, so an eigenvector (necessarily nonzero, 5.5) has $y \neq 0$ and hence $\lambda^2 = -3$, impossible for $\lambda \in \mathbf{R}$.

Problem 5A.6. Define $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(w, z) = (z, w)$. Find all eigenvalues and eigenvectors of T.

Solution. The eigenvalues are 1 and -1, with eigenvectors the nonzero vectors (w, w) and $(w, -w)$ respectively.
The equation $T(w, z) = \lambda(w, z)$ reads

$$z = \lambda w, \qquad w = \lambda z,$$

whence $(\lambda^2 - 1)w = 0$. If $w = 0$ then $z = \lambda w = 0$, excluded for an eigenvector, so $\lambda^2 = 1$, leaving $\lambda = \pm 1$.
(i) $\lambda = 1$: both equations say $z = w$, and $T(w, w) = (w, w)$, so

$$\operatorname{null}(T - I) = \operatorname{span}\big((1, 1)\big).$$

(ii) $\lambda = -1$: both say $z = -w$, and $T(w, -w) = -(w, -w)$, so

$$\operatorname{null}(T + I) = \operatorname{span}\big((1, -1)\big).$$

Each null space contains a nonzero vector, so both ± 1 really are eigenvalues.

Problem 5A.7. Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$. Find all eigenvalues and eigenvectors of T.

Solution. The eigenvalues are 0 and 5, with eigenvectors the nonzero vectors $(z_1, 0, 0)$ and $(0, 0, z_3)$ respectively.
The equation $T(z_1, z_2, z_3) = \lambda(z_1, z_2, z_3)$ reads

$$2z_2 = \lambda z_1, \qquad 0 = \lambda z_2, \qquad 5z_3 = \lambda z_3.$$

(i) $\lambda = 0$: these say $z_2 = z_3 = 0$, so

$$\operatorname{null} T = \operatorname{span}\big((1, 0, 0)\big) \neq \{0\},$$

and 0 is an eigenvalue.
(ii) $\lambda \neq 0$: the second equation gives $z_2 = 0$, then the first gives $z_1 = 0$, so an eigenvector is $(0, 0, z_3)$ with $z_3 \neq 0$ and the third equation forces $\lambda = 5$. Conversely $T(0, 0, z_3) = 5(0, 0, z_3)$, so

$$\operatorname{null}(T - 5I) = \operatorname{span}\big((0, 0, 1)\big).$$

Problem 5A.8. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that if λ is an eigenvalue of P, then $\lambda = 0$ or $\lambda = 1$.

Solution. Apply P to the eigenvalue equation. Take $v \neq 0$ with $Pv = \lambda v$ (5.5); then

$$\lambda v = Pv = P^2 v = P(\lambda v) = \lambda^2 v,$$

so $\lambda(\lambda - 1)v = 0$. Since $v \neq 0$ we get $\lambda(\lambda - 1) = 0$, and $\mathbf{F}$ has no zero divisors, so $\lambda = 0$ or $\lambda = 1$.

Problem 5A.9. Define $T \colon \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Tp = p'$. Find all eigenvalues and eigenvectors of T.

Solution. The only eigenvalue is 0, whose eigenvectors are the nonzero constant polynomials.
Degree settles it. Suppose $p \neq 0$ has degree m and $p' = \lambda p$ with $\lambda \neq 0$. If $m = 0$ then $p' = 0$ while λp is the nonzero constant λa_0; if $m \geq 1$ then

$$p'(x) = a_1 + 2a_2 x + \cdots + m a_m x^{m-1}$$

has degree $m - 1$ (its leading coefficient $m a_m$ is nonzero, $\mathbf{R}$ having characteristic 0) whereas $\deg(\lambda p) = m$. Either way $p' \neq \lambda p$, so no nonzero λ is an eigenvalue.

For $\lambda = 0$ the condition $p' = 0$ says $a_1 = 2a_2 = \cdots = ma_m = 0$, i.e. p is constant; the constant 1 is a nonzero such p, so 0 is an eigenvalue.

Problem 5A.10. Define $T \in \mathcal{L}(\mathcal{P}_4(\mathbf{R}))$ by $(Tp)(x) = xp'(x)$ for all $x \in \mathbf{R}$. Find all eigenvalues and eigenvectors of T.

Solution. The eigenvalues are $0, 1, 2, 3, 4$, the eigenvectors for k being the nonzero multiples of x^k; this is immediate from $Tx^j = x \cdot jx^{j-1} = jx^j$.
For completeness, write $p = \sum_{j=0}^{4} a_j x^j$, so that $Tp = \sum_{j=0}^{4} j\, a_j x^j$. Since $1, x, x^2, x^3, x^4$ is a basis of $\mathcal{P}_4(\mathbf{R})$, the equation $Tp = \lambda p$ is equivalent to

$$(j - \lambda)a_j = 0 \qquad \text{for } j = 0, 1, 2, 3, 4.$$

(i) $\lambda \notin \{0, 1, 2, 3, 4\}$: every $a_j = 0$, so $p = 0$ and λ is not an eigenvalue.
(ii) $\lambda = k \in \{0, 1, 2, 3, 4\}$: $a_j = 0$ for $j \neq k$, so $p(x) = a_k x^k$.

Problem 5A.11. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\alpha \in \mathbf{F}$. Prove that there exists $\delta > 0$ such that $T - \lambda I$ is invertible for all $\lambda \in \mathbf{F}$ such that $0 < |\alpha - \lambda| < \delta$.

Solution. Take δ to be the distance from α to the nearest eigenvalue of T other than α (and $\delta = 1$ if there is none). By 5.12 the set

$$E = \{\mu \in \mathbf{F} : \mu \text{ is an eigenvalue of } T,\ \mu \neq \alpha\}$$

is finite, say $E = \{\lambda_1, \ldots, \lambda_m\}$, so

$$\delta = \min\{|\alpha - \lambda_1|, \ldots, |\alpha - \lambda_m|\} > 0,$$

each $|\alpha - \lambda_j|$ being positive.
If $0 < |\alpha - \lambda| < \delta$, then $\lambda \neq \alpha$ and $\lambda \notin E$ (otherwise $|\alpha - \lambda| \geq \delta$), so λ is not an eigenvalue of T. As V is finite-dimensional, 5.7 then gives $T - \lambda I$ invertible.

Problem 5A.12. Suppose $V = U \oplus W$, where U and W are nonzero subspaces of V. Define $P \in \mathcal{L}(V)$ by $P(u + w) = u$ for each $u \in U$ and each $w \in W$. Find all eigenvalues and eigenvectors of P.

Solution. The eigenvalues are 0 and 1, with eigenvectors the nonzero vectors of W and of U respectively; both U and W are nonzero, so both are genuinely eigenvalues.
Since $Pu = u$ for $u \in U$, we have $P^2 = P$, so Exercise 5A.8 confines the eigenvalues to $\{0, 1\}$. Now decompose $v = u + w$ with $u \in U$, $w \in W$ (uniquely, by 1.41), so that $Pv = u$.
(i) $Pv = v$ reads $u = u + w$, i.e. $w = 0$; thus $\operatorname{null}(P - I) = U$.
(ii) $Pv = 0$ reads $u = 0$; thus $\operatorname{null} P = W$.

Problem 5A.13. Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible.

- (a) Prove that T and $S^{-1}TS$ have the same eigenvalues.
- (b) What is the relationship between the eigenvectors of T and the eigenvectors of $S^{-1}TS$?

Solution. (a) Conjugation respects the shift by λI: for each $\lambda \in \mathbf{F}$,

$$S^{-1}TS - \lambda I = S^{-1}(T - \lambda I)S,$$

since $S^{-1}(\lambda I)S = \lambda I$. As S and S^{-1} are bijective, the left side is injective exactly when $T - \lambda I$ is, and λ is an eigenvalue of an operator exactly when that operator minus λI fails to be injective (5.5). Hence T and $S^{-1}TS$ have the same eigenvalues.

(b) S carries eigenvectors of $S^{-1}TS$ to eigenvectors of T for the same eigenvalue. Indeed, for $w \neq 0$ (so $Sw \neq 0$, S being injective), applying the invertible S turns $(S^{-1}TS)w = \lambda w$ into the equivalent equation

$$T(Sw) = \lambda(Sw).$$

Equivalently, $\operatorname{null}(S^{-1}TS - \lambda I) = S^{-1}\bigl(\operatorname{null}(T - \lambda I)\bigr)$ for every λ.

Problem 5A.14. Give an example of an operator on $\mathbf{R}^4$ that has no (real) eigenvalues.

Solution. Take the quarter turn in each coordinate plane:

$$T(x_1, x_2, x_3, x_4) = (-x_2,\ x_1,\ -x_4,\ x_3),$$

which is linear since each output coordinate is linear in the input. If $Tv = \lambda v$ with $\lambda \in \mathbf{R}$, comparing coordinates gives

$$-x_2 = \lambda x_1, \qquad x_1 = \lambda x_2, \qquad -x_4 = \lambda x_3, \qquad x_3 = \lambda x_4.$$

Substituting the second into the first yields $(\lambda^2 + 1)x_2 = 0$, and $\lambda^2 + 1 > 0$ for real λ, so $x_2 = 0$ and then $x_1 = \lambda x_2 = 0$; the last two equations give $x_4 = x_3 = 0$ the same way. Thus $v = 0$, so T has no eigenvalue (5.5).

Problem 5A.15. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Show that λ is an eigenvalue of T if and only if λ is an eigenvalue of the dual operator $T' \in \mathcal{L}(V')$.

Solution. Dualizing turns surjectivity of $T - \lambda I$ into injectivity of $T' - \lambda I_{V'}$. Since $I'(\varphi) = \varphi \circ I = \varphi$, additivity and homogeneity of the dual map (3.120) give

$$(T - \lambda I)' = T' - \lambda I' = T' - \lambda I_{V'}.$$

Also $\dim V' = \dim V < \infty$ (3.111), so 5.7 applies to both T and T':

$$\begin{aligned}
\lambda \text{ is an eigenvalue of } T &\iff T - \lambda I \text{ is not surjective}\\
&\iff (T - \lambda I)' \text{ is not injective}\\
&\iff T' - \lambda I_{V'} \text{ is not injective}\\
&\iff \lambda \text{ is an eigenvalue of } T',
\end{aligned}$$

the first and last steps by 5.7 and the second by 3.129 (which needs V finite-dimensional), applied to $T - \lambda I$.

Problem 5A.16. Suppose $v_1, \ldots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda| \leq n \max\bigl\{ \bigl|\mathcal{M}(T)_{j,k}\bigr| : 1 \leq j, k \leq n \bigr\},$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \ldots, v_n$.
See Exercise 19 in Section 6A for a different bound on $|\lambda|$.

Solution. Test the eigenvector equation in the coordinate of largest modulus. Write $A = \mathcal{M}(T)$, so $Tv_k = \sum_j A_{j,k} v_j$ (3.31), and put $C = \max_{j,k} |A_{j,k}|$. Let $v = \sum_k c_k v_k$ be an eigenvector for λ, and choose m with $|c_m| = \max_k |c_k|$; since $v \neq 0$ we have $|c_m| > 0$. Expanding $Tv = \lambda v$ in the basis and comparing the (unique) coefficients of v_j,

$$\lambda c_j = \sum_{k=1}^{n} A_{j,k} c_k \qquad \text{for } j = 1, \ldots, n.$$

Taking $j = m$, the triangle inequality with $|A_{m,k}| \le C$ and $|c_k| \le |c_m|$ gives

$$|\lambda|\,|c_m| = \Big|\sum_{k=1}^{n} A_{m,k}c_k\Big| \le \sum_{k=1}^{n} |A_{m,k}|\,|c_k| \le nC|c_m|,$$

and dividing by $|c_m| > 0$ yields $|\lambda| \le nC$.

Problem 5A.17. Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{R}$. Prove that λ is an eigenvalue of T if and only if λ is an eigenvalue of the complexification $T_{\mathbf{C}}$.
See Exercise 33 in Section 3B for the definition of $T_{\mathbf{C}}$.

Solution. The eigenvector equation splits into its two components. Since λ is real, $\lambda(u+iv) = \lambda u + i\lambda v$, while $T_{\mathbf{C}}(u + iv) = Tu + iTv$; as elements of $V_{\mathbf{C}} = V \times V$ are equal exactly when both components agree,

$$T_{\mathbf{C}}(u + iv) = \lambda(u + iv) \iff Tu = \lambda u \text{ and } Tv = \lambda v.$$

(i) If $Tv = \lambda v$ with $v \neq 0$, then $v + i0 \neq 0$ in $V_{\mathbf{C}}$ and the equivalence gives $T_{\mathbf{C}}(v + i0) = \lambda(v + i0)$, so λ is an eigenvalue of $T_{\mathbf{C}}$.
(ii) If $T_{\mathbf{C}}(u + iv) = \lambda(u + iv)$ with $u + iv \neq 0$, then u, v are not both 0, and whichever is nonzero is an eigenvector of T corresponding to λ.

Problem 5A.18. Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{C}$. Prove that λ is an eigenvalue of the complexification $T_{\mathbf{C}}$ if and only if $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$.

Solution. Conjugation $\sigma(u + iv) = u - iv$ carries eigenvectors for λ to eigenvectors for $\overline{\lambda}$. Since the two directions are symmetric ($\sigma \circ \sigma$ is the identity, and $\overline{\overline{\lambda}} = \lambda$), it suffices to treat one. Three properties of σ, all read off from $T_{\mathbf{C}}(u + iv) = Tu + iTv$ and $(a + bi)(u + iv) = (au - bv) + i(av + bu)$:
(i) σ is injective, being the map $(u, v) \mapsto (u, -v)$ on $V_{\mathbf{C}} = V \times V$; in particular $\sigma(w) = 0$ only for $w = 0$.
(ii) σ is conjugate-homogeneous: with $\mu = a + bi$, both $\sigma(\mu(u + iv))$ and $\overline{\mu}\,\sigma(u + iv)$ equal $(au - bv) + i(-av - bu)$. (Check!)
(iii) σ commutes with $T_{\mathbf{C}}$, since $T(-v) = -Tv$:

$$\sigma\big(T_{\mathbf{C}}(u + iv)\big) = Tu + i\,T(-v) = T_{\mathbf{C}}\big(\sigma(u + iv)\big).$$

So if $T_{\mathbf{C}}w = \lambda w$ with $w \neq 0$, then $\sigma(w) \neq 0$ and

$$T_{\mathbf{C}}\big(\sigma(w)\big) = \sigma(T_{\mathbf{C}}w) = \sigma(\lambda w) = \overline{\lambda}\,\sigma(w),$$

so $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$.

Problem 5A.19. Show that the forward shift operator $T \in \mathcal{L}(\mathbf{F}^{\infty})$ defined by

$$T(z_1, z_2, \dots) = (0, z_1, z_2, \dots)$$

has no eigenvalues.

Solution. For every λ, the equation $Tz = \lambda z$ forces $z = 0$. Comparing coordinates in

$$(0, z_1, z_2, \dots) = (\lambda z_1, \lambda z_2, \lambda z_3, \dots)$$

gives $\lambda z_1 = 0$ and $z_k = \lambda z_{k+1}$ for every $k \ge 1$.
(i) $\lambda = 0$: the second family reads $z_k = 0$ for every k.
(ii) $\lambda \neq 0$: then $z_1 = 0$, and $z_{k+1} = z_k/\lambda$ propagates this up the sequence by induction, so again every $z_k = 0$.

An eigenvector must be nonzero, so T has no eigenvalues.

Problem 5A.20. Define the backward shift operator $S \in \mathcal{L}(\mathbf{F}^\infty)$ by

$$S(z_1, z_2, z_3, \dots) = (z_2, z_3, \dots).$$

(a) Show that every element of $\mathbf{F}$ is an eigenvalue of S.
(b) Find all eigenvectors of S.

Solution. (a) For $\lambda \in \mathbf{F}$ take the geometric sequence

$$w_\lambda = (1, \lambda, \lambda^2, \lambda^3, \dots) \qquad (\lambda^0 = 1),$$

which is nonzero (first coordinate 1) and satisfies $Sw_\lambda = (\lambda, \lambda^2, \dots) = \lambda w_\lambda$, the k-th coordinates being $\lambda^k = \lambda \cdot \lambda^{k-1}$. So every $\lambda \in \mathbf{F}$ is an eigenvalue.
(b) The eigenvectors are exactly the vectors $c\,w_\lambda$ with $\lambda \in \mathbf{F}$ and $c \neq 0$, such a vector corresponding to the eigenvalue λ. Each is nonzero and satisfies $S(cw_\lambda) = \lambda(cw_\lambda)$ by (a). Conversely, if $Sz = \lambda z$ with $z \neq 0$, comparing k-th coordinates in $(z_2, z_3, \dots) = (\lambda z_1, \lambda z_2, \dots)$ gives $z_{k+1} = \lambda z_k$, hence by induction

$$z_k = \lambda^{k-1} z_1 \qquad \text{for every positive integer } k.$$

Here $z_1 \neq 0$, since $z_1 = 0$ would force $z = 0$; so $z = z_1 w_\lambda$.

Problem 5A.21. Suppose $T \in \mathcal{L}(V)$ is invertible.
(a) Suppose $\lambda \in \mathbf{F}$ with $\lambda \neq 0$. Prove that λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.
(b) Prove that T and T^{-1} have the same eigenvectors.

Solution. Both parts rest on one computation: if $\lambda \neq 0$ and $Tv = \lambda v$, then applying T^{-1} gives $v = T^{-1}(\lambda v) = \lambda T^{-1} v$, so

$$T^{-1} v = \frac{1}{\lambda}\, v,$$

with the very same vector v. Since T^{-1} is invertible with $(T^{-1})^{-1} = T$, the same computation run with $T^{-1}, \frac{1}{\lambda}$ in place of T, λ turns $T^{-1}v = \frac{1}{\lambda} v$ back into $Tv = \lambda v$.
(a) If $\lambda \neq 0$ is an eigenvalue of T with eigenvector v, the display exhibits the nonzero v as an eigenvector of T^{-1} for $\frac{1}{\lambda}$; the reversed computation gives the converse.
(b) If v is an eigenvector of T, say $Tv = \lambda v$ with $v \neq 0$, then $\lambda \neq 0$ (an invertible T is injective, so $\operatorname{null} T = \{0\}$), and the display makes v an eigenvector of T^{-1}. Applying this to T^{-1}, whose inverse is T, gives the reverse inclusion. Hence T and T^{-1} have the same eigenvectors.

Problem 5A.22. Suppose $T \in \mathcal{L}(V)$ and there exist nonzero vectors u and w in V such that

$$Tu = 3w \quad \text{and} \quad Tw = 3u.$$

Prove that 3 or -3 is an eigenvalue of T.

Solution. The candidate eigenvectors are $u + w$ and $u - w$:

$$\begin{aligned} T(u + w) &= Tu + Tw = 3w + 3u = 3(u + w), \\ T(u - w) &= Tu - Tw = 3w - 3u = -3(u - w). \end{aligned}$$

They are not both zero, since $(u + w) + (u - w) = 2u \neq 0$. So:
(i) if $u + w \neq 0$, the first line makes 3 an eigenvalue of T;
(ii) if $u + w = 0$, then $u - w = 2u \neq 0$ and the second line makes -3 an eigenvalue of T.

Problem 5A.23. Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues.

Solution. By symmetry in S and T, it suffices to show every eigenvalue λ of ST is an eigenvalue of TS. Fix $v \neq 0$ with $(ST)v = \lambda v$.
(i) $\lambda \neq 0$. Then $w := Tv \neq 0$, since $Tv = 0$ would give $\lambda v = S0 = 0$ and hence $v = 0$; and

$$(TS)w = T\big((ST)v\big) = T(\lambda v) = \lambda Tv = \lambda w.$$

(ii) $\lambda = 0$. Then ST is not invertible by 5.7 (V finite-dimensional), so S or T is not invertible. Were TS invertible, S would be injective ($Sv = 0$ forces $(TS)v = 0$, so $v = 0$) hence invertible by 3.65, and then $T = (TS)S^{-1}$ would be invertible too, a contradiction. So TS is not invertible, and 0 is an eigenvalue of TS by 5.7.

Problem 5A.24. Suppose A is an n-by-n matrix with entries in $\mathbf{F}$. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $Tx = Ax$, where elements of $\mathbf{F}^n$ are thought of as n-by-1 column vectors.
(a) Suppose the sum of the entries in each row of A equals 1. Prove that 1 is an eigenvalue of T.
(b) Suppose the sum of the entries in each column of A equals 1. Prove that 1 is an eigenvalue of T.

Solution. (a) The eigenvector is $x = (1, \dots, 1)^{\mathsf{T}} \neq 0$: its image has j^{th} entry $\sum_{k=1}^{n} A_{j,k} \cdot 1 = 1$ by the row hypothesis, so $Tx = x = 1 \cdot x$.
(b) Here $T - I$ fails to be surjective. Put $B = A - I_n$, so $(T - I)x = Bx$; each column of B sums to 0, since subtracting I_n removes exactly 1 from the diagonal entry of column k. Hence for every $x \in \mathbf{F}^n$, swapping the order of summation,

$$\sum_{j=1}^{n} (Bx)_j = \sum_{k=1}^{n} x_k \Big(\sum_{j=1}^{n} B_{j,k} \Big) = 0,$$

so $\operatorname{range}(T - I) \subseteq U := \{y : y_1 + \cdots + y_n = 0\}$. Now U is the null space of the surjective linear functional $y \mapsto y_1 + \cdots + y_n$, so $\dim U = n - 1$ by 3.21 and $U \neq \mathbf{F}^n$. Thus $T - I$ is not surjective, and 1 is an eigenvalue of T by 5.7.

Problem 5A.25. Suppose $T \in \mathcal{L}(V)$ and u, w are eigenvectors of T such that $u + w$ is also an eigenvector of T. Prove that u and w are eigenvectors of T corresponding to the same eigenvalue.

Solution. Let $Tu = \alpha u$, $Tw = \beta w$, $T(u + w) = \lambda(u + w)$ with $u, w, u + w$ nonzero; we show $\alpha = \beta$. Linearity gives

$$\lambda u + \lambda w = Tu + Tw = \alpha u + \beta w,$$

that is, $(\alpha - \lambda)u + (\beta - \lambda)w = 0$. If $\alpha \neq \beta$, then u, w is linearly independent by 5.11 (eigenvectors for distinct eigenvalues), so both coefficients vanish and $\alpha = \lambda = \beta$ – contradicting $\alpha \neq \beta$. Hence $\alpha = \beta$.

Problem 5A.26. Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector in V is an eigenvector of T. Prove that T is a scalar multiple of the identity operator.

Solution. If $V = \{0\}$ then $T = 0 = 0 \cdot I$, so assume $V \neq \{0\}$. By hypothesis each nonzero $v \in V$ has a scalar a_v with $Tv = a_v v$, and a_v is unique because $v \neq 0$. It suffices to show $a_u = a_w$ for all nonzero u, w.
(i) u, w linearly dependent. Then $w = cu$ with $c \neq 0$ (as $u, w \neq 0$), so

$$a_w w = Tw = c\,Tu = c\,a_u u = a_u w,$$

whence $a_w = a_u$.

(ii) u, w linearly independent. Then $u + w \neq 0$ and

$$a_{u+w}u + a_{u+w}w = T(u+w) = a_u u + a_w w,$$

so independence forces $a_u = a_{u+w} = a_w$.
Writing λ for the common value, $Tv = \lambda v$ for all nonzero v, and $T0 = 0 = \lambda \cdot 0$. Hence $T = \lambda I$.

Problem 5A.27. Suppose that V is finite-dimensional and $k \in \{1, \ldots, \dim V - 1\}$. Suppose $T \in \mathcal{L}(V)$ is such that every subspace of V of dimension k is invariant under T. Prove that T is a scalar multiple of the identity operator.

Solution. It suffices to show every nonzero $v \in V$ is an eigenvector of T, for then 5A.26 gives $T = \lambda I$. Set $n = \dim V$, so $n \geq 2$ because $k \in \{1, \ldots, n-1\}$.
Suppose instead $v \neq 0$ with $Tv \notin \operatorname{span}(v)$, so that v, Tv is linearly independent; extend it by 2.32 to a basis $v, Tv, w_3, \ldots, w_n$ of V and put

$$U = \operatorname{span}(v, w_3, \ldots, w_{k+1}),$$

legitimate since $k + 1 \leq n$ (the list $w_3, \ldots, w_{k+1}$ is empty when $k = 1$). Being a sublist of a basis, $v, w_3, \ldots, w_{k+1}$ is linearly independent of length k, so $\dim U = k$ and U is invariant under T by hypothesis. Then $Tv \in U$ expresses Tv in the basis $v, Tv, w_3, \ldots, w_n$ with coefficient 0 on Tv, while the obvious expression has coefficient 1; uniqueness of basis representations (2.28) gives $1 = 0$, a contradiction.
Hence $Tv \in \operatorname{span}(v)$ for every nonzero v, which is exactly the hypothesis of 5A.26. So T is a scalar multiple of the identity.

Problem 5A.28. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has at most $1 + \dim \operatorname{range} T$ distinct eigenvalues.

Solution. Let $\lambda_1, \ldots, \lambda_m$ be distinct eigenvalues of T with eigenvectors $v_1, \ldots, v_m$; this list is linearly independent by 5.11. At most one λ_j is 0, so relabel to make $\lambda_1, \ldots, \lambda_{m-1}$ nonzero. Then for $j \leq m-1$,

$$v_j = \frac{1}{\lambda_j} T v_j = T\Big(\frac{v_j}{\lambda_j}\Big) \in \operatorname{range} T,$$

so $v_1, \ldots, v_{m-1}$ is a linearly independent list inside $\operatorname{range} T$, a subspace of the finite-dimensional V. Comparing it with a basis of $\operatorname{range} T$, which spans, 2.22 gives $m - 1 \leq \dim \operatorname{range} T$, that is, $m \leq 1 + \dim \operatorname{range} T$.

Problem 5A.29. Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ and -4, 5, and $\sqrt{7}$ are eigenvalues of T. Prove that there exists $x \in \mathbf{R}^3$ such that $Tx - 9x = (-4, 5, \sqrt{7})$.

Solution. It is enough that $T - 9I$ be surjective. Since $-4, 5, \sqrt{7}$ are three distinct eigenvalues and $\dim \mathbf{R}^3 = 3$, by 5.12 the operator T has no other eigenvalue; as $9 \notin \{-4, 5, \sqrt{7}\}$ (note $\sqrt{7} < 3$), the number 9 is not an eigenvalue of T. So $T - 9I$ is injective by 5.7, hence surjective by 3.65. Thus some $x \in \mathbf{R}^3$ has $(T - 9I)x = (-4, 5, \sqrt{7})$.

Problem 5A.30. Suppose $T \in \mathcal{L}(V)$ and $(T - 2I)(T - 3I)(T - 4I) = 0$. Suppose λ is an eigenvalue of T. Prove that $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

Solution. Take $v \neq 0$ with $Tv = \lambda v$. Since $(T - \mu I)v = (\lambda - \mu)v$ for every $\mu \in \mathbf{F}$, applying the three factors to v from the right gives

$$(T-4I)v = (\lambda-4)v,$$
$$(T-3I)(T-4I)v = (\lambda-4)(\lambda-3)v,$$
$$(T-2I)(T-3I)(T-4I)v = (\lambda-2)(\lambda-3)(\lambda-4)v.$$

The left side of the last line is 0 by hypothesis, and $v \neq 0$, so $(\lambda-2)(\lambda-3)(\lambda-4) = 0$. As $\mathbf{F}$ has no zero divisors, $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

Problem 5A.31. Give an example of $T \in \mathcal{L}(\mathbf{R}^2)$ such that $T^4 = -I$.

Solution. Take T to be counterclockwise rotation by 45 degrees, that is,

$$T(x,y) = \left(\frac{x-y}{\sqrt{2}}, \frac{x+y}{\sqrt{2}}\right),$$

which is linear (Check!). Squaring,

$$T^2(x,y) = \left(\frac{-2y}{2}, \frac{2x}{2}\right) = (-y,x),$$

so $T^4(x,y) = T^2(-y,x) = (-x,-y)$. Hence $T^4 = -I$.

Problem 5A.32. Suppose $T \in \mathcal{L}(V)$ has no eigenvalues and $T^4 = I$. Prove that $T^2 = -I$.

Solution. Because polynomials in T factor as polynomials do (5.17) and $z^4-1 = (z-1)(z+1)(z^2+1)$, the hypothesis $T^4 = I$ gives

$$(T-I)(T+I)(T^2+I) = T^4 - I = 0.$$

Suppose $T^2+I \neq 0$, say $w := (T^2+I)v \neq 0$; then $(T-I)(T+I)w = 0$. Put $u = (T+I)w$.
(i) $u = 0$. Then $Tw = -w$ with $w \neq 0$, so -1 is an eigenvalue of T.
(ii) $u \neq 0$. Then $(T-I)u = 0$, so $Tu = u$ and 1 is an eigenvalue of T.
Both contradict the hypothesis that T has no eigenvalues. Hence $T^2+I = 0$, that is, $T^2 = -I$.

Problem 5A.33. Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.
(a) Prove that T is injective if and only if T^m is injective.
(b) Prove that T is surjective if and only if T^m is surjective.

Solution. Both directions are direct, since V is not assumed finite-dimensional; throughout $T^m = T^{m-1}T = T\,T^{m-1}$ (5.13).
(a) By 3.15, injectivity means null space $\{0\}$. If T is injective, induct: assuming T^{m-1} injective, $T^m v = T^{m-1}(Tv) = 0$ forces $Tv = 0$, hence $v = 0$. Conversely, if T^m is injective and $Tv = 0$, then

$$T^m v = T^{m-1}(Tv) = 0,$$

so $v = 0$ (for $m = 1$ this reads $Tv = 0$ directly).
(b) Surjectivity means $T(V) = V$. If T is surjective, induct: $T^m(V) = T^{m-1}(T(V)) = T^{m-1}(V) = V$. Conversely, if $T^m(V) = V$, then

$$V = T^m(V) = T\big(T^{m-1}(V)\big) \subseteq T(V) \subseteq V,$$

the first inclusion because $T^{m-1}(V) \subseteq V$. Hence $T(V) = V$.

Problem 5A.34. Suppose V is finite-dimensional and $v_1, \ldots, v_m \in V$. Prove that the list $v_1, \ldots, v_m$ is linearly independent if and only if there exists $T \in \mathcal{L}(V)$ such that $v_1, \ldots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues.

Solution. One direction is immediate from 5.11: eigenvectors corresponding to distinct eigenvalues form a linearly independent list.
For the other, suppose $v_1, \ldots, v_m$ is linearly independent and extend it by 2.32 to a basis $v_1, \ldots, v_n$ of V, where $n = \dim V$. The linear map lemma 3.4, applied to this basis, supplies $T \in \mathcal{L}(V)$ with

$$Tv_k = k\, v_k \qquad \text{for } k = 1, \ldots, n.$$

Each v_k is nonzero (a linearly independent list contains no 0), so $v_1, \ldots, v_m$ are eigenvectors of T for the eigenvalues $1, \ldots, m$, which are distinct in $\mathbf{F}$.

Problem 5A.35. Suppose that $\lambda_1, \ldots, \lambda_n$ is a list of distinct real numbers. Prove that the list $e^{\lambda_1 x}, \ldots, e^{\lambda_n x}$ is linearly independent in the vector space of real-valued functions on $\mathbf{R}$.
Hint: Let $V = \operatorname{span}(e^{\lambda_1 x}, \ldots, e^{\lambda_n x})$, and define an operator $D \in \mathcal{L}(V)$ by $Df = f'$. Find eigenvalues and eigenvectors of D.

Solution. Write f_k for the function $x \mapsto e^{\lambda_k x}$ and set $V = \operatorname{span}(f_1, \ldots, f_n)$, a subspace of the space of real-valued functions on $\mathbf{R}$. Differentiation $Df = f'$ is linear and, since

$$f_k' = \lambda_k f_k \in V \qquad \text{for } k = 1, \ldots, n,$$

it maps each spanning vector of V into V, hence maps V into V; so $D \in \mathcal{L}(V)$. Each f_k is nonzero in V because $f_k(0) = 1$, so the display makes f_k an eigenvector of D for the eigenvalue λ_k, and the λ_k are distinct by hypothesis. By 5.11 the list $f_1, \ldots, f_n$ is linearly independent in V, hence in the larger space, linear independence being a statement about vanishing linear combinations only.

Problem 5A.36. Suppose that $\lambda_1, \ldots, \lambda_n$ is a list of distinct positive numbers. Prove that the list $\cos(\lambda_1 x), \ldots, \cos(\lambda_n x)$ is linearly independent in the vector space of real-valued functions on $\mathbb{R}$.

Solution. Run the device of 5A.35 with the second derivative. Write c_k for $x \mapsto \cos(\lambda_k x)$, set $V = \operatorname{span}(c_1, \ldots, c_n)$, and define $Tf = f''$, which is linear and satisfies

$$Tc_k = -\lambda_k^2\, c_k \in V \qquad \text{for } k = 1, \ldots, n,$$

so T maps a spanning list of V into V and hence $T \in \mathcal{L}(V)$. Each c_k is nonzero since $c_k(0) = 1$, so c_k is an eigenvector for the eigenvalue $-\lambda_k^2$. These eigenvalues are distinct: $\lambda_j^2 = \lambda_k^2$ gives $(\lambda_j - \lambda_k)(\lambda_j + \lambda_k) = 0$, and $\lambda_j + \lambda_k > 0$ by positivity, so $\lambda_j = \lambda_k$ and $j = k$. By 5.11 the list is linearly independent in V, hence in the space of all real-valued functions on $\mathbb{R}$.

Problem 5A.37. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(S) = TS$$

for each $S \in \mathcal{L}(V)$. Prove that the set of eigenvalues of T equals the set of eigenvalues of $\mathcal{A}$.

Solution. If $V = \{0\}$ then $\mathcal{L}(V) = \{0\}$ too and both operators have no eigenvalues, so assume $n = \dim V \geq 1$.
If λ is an eigenvalue of $\mathcal{A}$, choose $S \neq 0$ with $TS = \lambda S$ and v with $Sv \neq 0$; applying the equation to v gives $T(Sv) = \lambda(Sv)$, so λ is an eigenvalue of T.
Conversely, let $Tv = \lambda v$ with $v \neq 0$, fix a basis $v_1, \ldots, v_n$ of V, and use 3.4 on that basis to get

$S \in \mathcal{L}(V)$ with

$$Sv_1 = v, \qquad Sv_j = 0 \text{ for } j = 2, \ldots, n,$$

so $S \neq 0$. For $w = a_1 v_1 + \cdots + a_n v_n$ we have $Sw = a_1 v$, whence

$$(TS)w = T(a_1 v) = \lambda(a_1 v) = \lambda(Sw).$$

Thus $\mathcal{A}(S) = TS = \lambda S$ with $S \neq 0$, making λ an eigenvalue of $\mathcal{A}$.

Problem 5A.38. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V invariant under T. The quotient operator $T/U \in \mathcal{L}(V/U)$ is defined by

$$(T/U)(v + U) = Tv + U$$

for each $v \in V$.
(a) Show that the definition of T/U makes sense (which requires using the condition that U is invariant under T) and show that T/U is an operator on V/U.
(b) Show that each eigenvalue of T/U is an eigenvalue of T.

Solution. (a) Independence of the representative is exactly where invariance enters: if $v + U = w + U$, then $v - w \in U$, so $T(v - w) \in U$ by invariance, that is, $Tv - Tw \in U$, which says

$$Tv + U = Tw + U.$$

So T/U is a well-defined function on V/U, and it is linear because

$$(T/U)\big((v + U) + (w + U)\big) = T(v + w) + U = (Tv + U) + (Tw + U),$$
$$(T/U)\big(\alpha(v + U)\big) = T(\alpha v) + U = \alpha\big((T/U)(v + U)\big).$$

Hence $T/U \in \mathcal{L}(V/U)$.
(b) Let λ be an eigenvalue of T/U, so some $v \notin U$ has $Tv + U = \lambda v + U$, i.e. $u := (T - \lambda I)v \in U$. Since U is invariant under T and under I, it is invariant under $T - \lambda I$, so $(T - \lambda I)|_U \in \mathcal{L}(U)$ with U finite-dimensional.
(i) $(T - \lambda I)|_U$ not injective. Some nonzero $u' \in U$ has $Tu' = \lambda u'$.
(ii) $(T - \lambda I)|_U$ injective, hence surjective onto U by 3.65. Pick $u' \in U$ with $(T - \lambda I)u' = u$; then

$$(T - \lambda I)(v - u') = u - u = 0,$$

and $v - u' \neq 0$ since $v \notin U$.
Either way λ is an eigenvalue of T.

Problem 5A.39. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an eigenvalue if and only if there exists a subspace of V of dimension $\dim V - 1$ that is invariant under T.

Solution. For $V = \{0\}$ both sides fail (no nonzero vectors; no subspace of dimension -1), and in each direction below the assumed hypothesis forces $\dim V \geq 1$.
Suppose T has an eigenvalue λ and put $W = \operatorname{range}(T - \lambda I)$. Then $T - \lambda I$ is not injective by 5.7, hence not surjective by 3.65, so $W \neq V$ and $\dim W \leq n - 1$ where $n = \dim V$. Extend a basis $w_1, \ldots, w_m$ of W to a basis $w_1, \ldots, w_m, u_{m+1}, \ldots, u_n$ of V (2.32) and let U be the span of its first $n - 1$ vectors, so $\dim U = n - 1$ and $W \subseteq U$ (as $m \leq n - 1$). Then U is invariant:

$$Tu = (T - \lambda I)u + \lambda u \in W + U \subseteq U \qquad (u \in U).$$

Conversely, let U be invariant under T with $\dim U = \dim V - 1$, so $\dim V/U = 1$ by 3.105 and $T/U \in \mathcal{L}(V/U)$ by 5A.38(a). Any nonzero $w \in V/U$ spans V/U, so $(T/U)w = \lambda w$ for some $\lambda \in \mathbf{F}$; thus λ is an eigenvalue of T/U, hence of T by 5A.38(b).

Problem 5A.40. Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial. Prove that

$$p(STS^{-1}) = S\,p(T)\,S^{-1}.$$

Solution. Everything follows from $(STS^{-1})^k = ST^kS^{-1}$ for $k \geq 0$. This holds for $k = 0$ (both sides are I, by the convention $R^0 = I$ of 5.13), and inductively

$$(STS^{-1})^{k+1} = (ST^kS^{-1})(STS^{-1}) = ST^k(S^{-1}S)TS^{-1} = ST^{k+1}S^{-1}.$$

Now write $p(z) = a_0 + a_1z + \cdots + a_mz^m$. By 5.14 and the previous display,

$$p(STS^{-1}) = a_0SIS^{-1} + a_1STS^{-1} + \cdots + a_mST^mS^{-1},$$

and $R \mapsto SRS^{-1}$ is linear (composition distributes over addition and commutes with scalars), so the sum collapses to $S\,p(T)\,S^{-1}$.

Problem 5A.41. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. Prove that U is invariant under $p(T)$ for every polynomial $p \in \mathcal{P}(\mathbf{F})$.

Solution. First, U is invariant under every power T^k: true for $k = 0$ since $T^0 = I$ (5.13), and if $T^ku \in U$ for all $u \in U$, then $T^{k+1}u = T(T^ku) \in U$ by invariance under T. So for $p(z) = a_0 + a_1z + \cdots + a_mz^m$ and $u \in U$, by 5.14

$$p(T)u = a_0u + a_1(Tu) + \cdots + a_m(T^mu) \in U,$$

each term lying in U and U being closed under scalar multiples and sums.

Problem 5A.42. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, \ldots, x_n) = (x_1, 2x_2, 3x_3, \ldots, nx_n)$.
(a) Find all eigenvalues and eigenvectors of T.
(b) Find all subspaces of $\mathbf{F}^n$ that are invariant under T.

Solution. With $e_1, \ldots, e_n$ the standard basis, the definition of T says $Te_j = j\,e_j$.
(a) The eigenvalues are exactly $1, 2, \ldots, n$, with $E(j, T) = \operatorname{span}(e_j)$; the eigenvectors for j are the nonzero multiples of e_j. Indeed, comparing coordinates, $Tx = \lambda x$ says

$$(k - \lambda)\,x_k = 0 \qquad \text{for } k = 1, \ldots, n.$$

If $\lambda \notin \{1, \ldots, n\}$ every factor $k - \lambda$ is nonzero, forcing $x = 0$; if $\lambda = j$, the equations with $k \neq j$ force $x_k = 0$, leaving $x \in \operatorname{span}(e_j)$.
(b) The invariant subspaces are exactly the 2^n coordinate subspaces

$$U_\Omega = \operatorname{span}(e_j : j \in \Omega), \qquad \Omega \subseteq \{1, \ldots, n\}.$$

Each is invariant, since $Te_j = je_j \in U_\Omega$ on a spanning list. Conversely, let U be invariant and set

$$p_j(z) = \prod_{k \neq j,\ 1 \leq k \leq n} \frac{z - k}{j - k} \in \mathcal{P}(\mathbf{F}),$$

legitimate as $j - k \neq 0$ in $\mathbf{F}$, with $p_j(j) = 1$ and $p_j(k) = 0$ for $k \neq j$. From $T^me_k = k^me_k$ and 5.14 we get $p(T)e_k = p(k)e_k$ for every polynomial p, so for $x = x_1e_1 + \cdots + x_ne_n \in U$, invariance of U under $p_j(T)$ (5A.41) gives

$$p_j(T)x = \sum_{k=1}^{n} x_k\,p_j(k)\,e_k = x_j\,e_j \in U.$$

Hence $x_j \neq 0$ implies $e_j \in U$. With $\Omega = \{j : e_j \in U\}$, every $x \in U$ has $x_j = 0$ off Ω, so $U \subseteq U_\Omega$, while $U_\Omega \subseteq U$ by definition of Ω. These 2^n subspaces are distinct, since $e_j \in U_\Omega$ if and only if $j \in \Omega$ by independence of $e_1, \ldots, e_n$.

Problem 5A.43. Suppose that V is finite-dimensional, $\dim V > 1$, and $T \in \mathcal{L}(V)$. Prove that

$$\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V).$$

Solution. The set $\mathcal{E} = \{p(T) : p \in \mathcal{P}(\mathbf{F})\}$ is commutative, since $p(T)q(T) = q(T)p(T)$ by 5.17(b), whereas $\mathcal{L}(V)$ is not. Indeed, let $n = \dim V \geq 2$, take a basis $v_1, \ldots, v_n$ of V (2.31), and use 3.4 to get $R, S \in \mathcal{L}(V)$ with

$$Rv_1 = v_2, \qquad Sv_1 = v_1, \qquad Rv_j = Sv_j = 0 \text{ for } j \geq 2.$$

Then

$$(RS)v_1 = Rv_1 = v_2 \neq 0 = Sv_2 = (SR)v_1,$$

so $RS \neq SR$. Were $\mathcal{E} = \mathcal{L}(V)$, these R, S would lie in $\mathcal{E}$ and hence commute. So $\mathcal{E} \neq \mathcal{L}(V)$.
Method (2): suppose $\mathcal{E} = \mathcal{L}(V)$. Taking $q(z) = z$ in 5.17(b), every element of $\mathcal{E}$, hence every operator on V, commutes with T. Then every nonzero v is an eigenvector of T: otherwise v, Tv is linearly independent, extends to a basis $v, Tv, u_3, \ldots, u_n$ (2.32), and 3.4 supplies S with $Sv = v$, $S(Tv) = 0$, $Su_j = 0$, giving $(ST)v = 0 \neq Tv = (TS)v$ (that $Tv \neq 0$ is part of the independence). So $T = \lambda I$ by 5A.26, whence $p(T) = p(\lambda)I$ for every p and $\dim \mathcal{E} = \dim \operatorname{span}(I) = 1$. But $\dim \mathcal{L}(V) = n^2 \geq 4$ by 3.72, a contradiction.

5.2 Exercises 5B

Problem 5B.1. Suppose $T \in \mathcal{L}(V)$. Prove that 9 is an eigenvalue of T^2 if and only if 3 or -3 is an eigenvalue of T.

Solution. If $Tv = \pm 3v$ with $v \neq 0$, then

$$T^2 v = \pm 3\, Tv = (\pm 3)^2 v = 9v,$$

so 9 is an eigenvalue of T^2.
Conversely, take $v \neq 0$ with $T^2 v = 9v$. Since polynomials in T factor as polynomials do (5.17), $T^2 - 9I = (T - 3I)(T + 3I)$, so with $w = (T + 3I)v$ we get $(T - 3I)w = 0$.
(i) $w = 0$: then $Tv = -3v$ with $v \neq 0$, so -3 is an eigenvalue of T.
(ii) $w \neq 0$: then $Tw = 3w$, so 3 is an eigenvalue of T.

Problem 5B.2. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ has no eigenvalues. Prove that every subspace of V invariant under T is either $\{0\}$ or infinite-dimensional.

Solution. Suppose some invariant U were nonzero and finite-dimensional. Invariance makes $T|_U \in \mathcal{L}(U)$, and U is a nonzero finite-dimensional complex vector space, so by 5.19 there is $u \in U$ with $u \neq 0$ and

$$Tu = (T|_U)u = \lambda u$$

for some $\lambda \in \mathbf{C}$. Then λ is an eigenvalue of T, contradicting the hypothesis.

Problem 5B.3. Suppose n is an integer with $n > 1$ and $T \in \mathcal{L}(\mathbf{F}^n)$ is defined by

$$T(x_1, \ldots, x_n) = (x_1 + \cdots + x_n, \ldots, x_1 + \cdots + x_n).$$

(a) Find all eigenvalues and eigenvectors of T.
(b) Find the minimal polynomial of T.
[The matrix of T with respect to the standard basis of $\mathbf{F}^n$ consists of all 1's.]

Solution. (a) The eigenvalues are 0 and n; the eigenvectors for 0 are the nonzero x with $x_1+\cdots+x_n=0$, and those for n are the nonzero multiples of $\mathbf{1}=(1,\ldots,1)$.
Writing $s(x)=x_1+\cdots+x_n$, the definition of T says $Tx=s(x)\mathbf{1}$. Hence $Tx=0$ iff $s(x)=0$, and since s is a surjective linear functional, 3.21 gives

$$\dim \operatorname{null} T = n-1 \geq 1,$$

so 0 is an eigenvalue with the stated eigenvectors. Also $T\mathbf{1}=n\mathbf{1}$ with $\mathbf{1}\neq 0$. If $Tx=\lambda x$ with $x\neq 0$ and $\lambda\neq 0$, then $x=\lambda^{-1}s(x)\mathbf{1}=c\mathbf{1}$ with $c\neq 0$, and applying T gives $\lambda c\mathbf{1}=cn\mathbf{1}$, so $\lambda=n$.
(b) The minimal polynomial is z^2-nz. For every $x\in\mathbf{F}^n$,

$$T^2x = s(x)\,T\mathbf{1} = n\,s(x)\,\mathbf{1} = n\,Tx,$$

so $T^2-nT=0$ and the minimal polynomial divides $z(z-n)$ by 5.29. By 5.27(a) its zeros are the eigenvalues 0 and n, which are distinct because $n>1$; hence its degree is at least 2, and being monic of degree 2 it equals z^2-nz.

Problem 5B.4. Suppose $\mathbf{F}=\mathbf{C}$, $T\in\mathcal{L}(V)$, $p\in\mathcal{P}(\mathbf{C})$ is a nonconstant polynomial, and $\alpha\in\mathbf{C}$. Prove that α is an eigenvalue of $p(T)$ if and only if $\alpha=p(\lambda)$ for some eigenvalue λ of T.

Solution. Suppose first $\alpha=p(\lambda)$ with $Tv=\lambda v$, $v\neq 0$. Then $T^kv=\lambda^kv$ for every $k\geq 0$, so writing $p(z)=a_0+a_1z+\cdots+a_mz^m$,

$$p(T)v=\sum_{k=0}^m a_kT^kv=\Big(\sum_{k=0}^m a_k\lambda^k\Big)v=p(\lambda)v=\alpha v,$$

and $v\neq 0$, so α is an eigenvalue of $p(T)$.
Conversely, suppose α is an eigenvalue of $p(T)$, and set $q(z)=p(z)-\alpha$. Then q is nonconstant, say $\deg q=\deg p=m\geq 1$, so 4.13 factors it as $q(z)=c(z-\lambda_1)\cdots(z-\lambda_m)$ with $c\neq 0$. Because $r\mapsto r(T)$ preserves products (5.17),

$$p(T)-\alpha I=q(T)=c\,(T-\lambda_1 I)\cdots(T-\lambda_m I).$$

The left side is not injective, and a composition of injective operators is injective, so $T-\lambda_jI$ fails to be injective for some j; that is, λ_j is an eigenvalue of T. Since $q(\lambda_j)=0$, we get $\alpha=p(\lambda_j)$.

Problem 5B.5. Give an example of an operator on $\mathbf{R}^2$ that shows the result in Exercise 4 does not hold if $\mathbf{C}$ is replaced with $\mathbf{R}$.

Solution. Take $T(x,y)=(-y,x)$ on $\mathbf{R}^2$, $p(z)=z^2$, and $\alpha=-1$.
Then $T^2(x,y)=T(-y,x)=-(x,y)$, so $p(T)=T^2=-I$ and $\alpha=-1$ is an eigenvalue of $p(T)$. But T has no eigenvalues: $T(x,y)=\lambda(x,y)$ says $-y=\lambda x$ and $x=\lambda y$, whence $(\lambda^2+1)y=0$, so $y=0$ and then $x=0$ (as $\lambda^2+1>0$ for real λ). Hence no eigenvalue λ of T satisfies $\alpha=p(\lambda)$, and Exercise 4 fails over $\mathbf{R}$.

Problem 5B.6. Suppose $T\in\mathcal{L}(\mathbf{F}^2)$ is defined by $T(w,z)=(-z,w)$. Find the minimal polynomial of T.

Solution. The minimal polynomial of T is z^2+1, for $\mathbf{F}=\mathbf{R}$ and for $\mathbf{F}=\mathbf{C}$. Indeed

$$T^2(w,z)=T(-z,w)=-(w,z),$$

so $T^2+I=0$ and the minimal polynomial divides z^2+1 by 5.29. It is not 1 (which gives $I\neq 0$) nor of the form $z-c$, since $T(1,0)=(0,1)$ is not a multiple of $(1,0)$, so $T\neq cI$. Hence its degree is 2 and it equals z^2+1.

Problem 5B.7. (a) Give an example of $S, T \in \mathcal{L}(\mathbf{F}^2)$ such that the minimal polynomial of ST does not equal the minimal polynomial of TS.
(b) Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that if at least one of S, T is invertible, then the minimal polynomial of ST equals the minimal polynomial of TS.
[Hint: Show that if S is invertible and $p \in \mathcal{P}(\mathbf{F})$, then $p(TS) = S^{-1}p(ST)S$.]

Solution. (a) Take $S(w,z) = (z,0)$ and $T(w,z) = (w,0)$. Then

$$(ST)(w,z) = S(w,0) = (0,0), \qquad (TS)(w,z) = T(z,0) = (z,0),$$

so $ST = 0$, with minimal polynomial z, while $TS \neq 0$ (it sends $(0,1)$ to $(1,0)$) satisfies $(TS)^2(w,z) = (TS)(z,0) = 0$. Thus z^2 annihilates TS and no degree-1 monic polynomial does ($TS = cI$ applied to $(1,0)$ forces $c = 0$, hence $TS = 0$), so the minimal polynomial of TS is $z^2 \neq z$.
(b) The statement is symmetric in S and T, so assume S is invertible. For every integer $k \geq 0$,

$$S^{-1}(ST)^k S = (S^{-1}S)(TS)^k = (TS)^k,$$

by regrouping the $2k$ alternating factors (the case $k = 0$ being $S^{-1}IS = I$). Hence, for $p(z) = a_0 + \cdots + a_m z^m$, linearity of $A \mapsto S^{-1}AS$ gives the hint:

$$p(TS) = \sum_{k=0}^{m} a_k S^{-1}(ST)^k S = S^{-1}p(ST)S.$$

Since S is invertible, $p(TS) = 0$ if and only if $p(ST) = 0$; so ST and TS are annihilated by exactly the same polynomials. As V is finite-dimensional, each has a minimal polynomial, namely the unique monic polynomial of smallest degree in that common set (5.22), so the two minimal polynomials coincide.

Problem 5B.8. Suppose $T \in \mathcal{L}(\mathbb{R}^2)$ is the operator of counterclockwise rotation by 1°. Find the minimal polynomial of T.
[Because $\dim \mathbb{R}^2 = 2$, the degree of the minimal polynomial of T is at most 2. Thus the minimal polynomial of T is not the tempting polynomial $x^{180} + 1$, even though $T^{180} = -I$.]

Solution. The minimal polynomial of T is

$$z^2 - \left(2\cos\tfrac{\pi}{180}\right) z + 1.$$

Write $\theta = \pi/180$. With respect to the standard basis of $\mathbb{R}^2$,

$$\mathcal{M}(T) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

which in the notation of Exercise 11 has $a = d = \cos\theta$, $b = \sin\theta$, and $c = -\sin\theta$. Here $b = \sin\frac{\pi}{180} \neq 0$, so Exercise 11(b) puts us in the second case and the minimal polynomial is

$$z^2 - (a+d)z + (ad - bc) = z^2 - (2\cos\theta)z + (\cos^2\theta + \sin^2\theta),$$

which is the polynomial displayed above.

Problem 5B.9. Suppose $T \in \mathcal{L}(V)$ is such that with respect to some basis of V, all entries of the matrix of T are rational numbers. Explain why all coefficients of the minimal polynomial of T are rational numbers.

Solution. Let $n = \dim V$ and let $A = \mathcal{M}(T)$ with respect to the given basis, so that A and all its powers have rational entries and, since $\mathcal{M}$ is injective, linear, and multiplicative (3.43), a polynomial

q satisfies $q(T) = 0$ if and only if $q(A) = 0$. Let p be the minimal polynomial of T and $m = \deg p$. The point is that $\mathbf{F}$ is $\mathbb{R}$ or $\mathbb{C}$, not $\mathbb{Q}$, so the following bridge is needed.
Lemma. If B is a rational matrix with k columns and $Bx = 0$ has a nonzero solution in $\mathbf{F}^k$, then it has one in $\mathbb{Q}^k$. Indeed, row reducing B uses only rational scalars, so its reduced row echelon form R is rational, and row operations are reversible and hence preserve the solution set over $\mathbf{F}$ and over $\mathbb{Q}$ alike. Were every column of R a pivot column, $Rx = 0$ would force $x = 0$ over $\mathbf{F}$; so some column is free, and setting that free variable to 1, the other free variables to 0, and solving for the pivot variables gives a nonzero solution in $\mathbb{Q}^k$.
The rational n-by-n matrices form a $\mathbb{Q}$-vector space of dimension n^2, so $I, A, \ldots, A^{n^2}$ are linearly dependent over $\mathbb{Q}$; let $m' \leq n^2$ be smallest with $I, A, \ldots, A^{m'}$ dependent over $\mathbb{Q}$, say

$$b_0 I + b_1 A + \cdots + b_{m'} A^{m'} = 0$$

with $b_j \in \mathbb{Q}$ not all 0. Minimality of m' forces $b_{m'} \neq 0$, so dividing by it produces a monic $q \in \mathcal{P}(\mathbb{Q})$ of degree m' with $q(A) = 0$ and hence $q(T) = 0$. By 5.29 the nonzero polynomial q is a polynomial multiple of p, so $m' \geq m$.
For the reverse inequality, $p(A) = 0$ is an $\mathbf{F}$-linear dependence among $I, A, \ldots, A^m$; listing the entries of these $m + 1$ rational matrices as the columns of a rational matrix B, it is a nonzero solution of $Bx = 0$ in $\mathbf{F}^{m+1}$, so the Lemma yields a rational dependence among $I, A, \ldots, A^m$ and thus $m' \leq m$. Hence $m' = m$, and q is monic of degree m with $q(T) = 0$, so $q = p$ by the uniqueness in 5.22. All coefficients of p are therefore rational.

Problem 5B.10. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$. Prove that

$$\operatorname{span}(v, Tv, \ldots, T^m v) = \operatorname{span}(v, Tv, \ldots, T^{\dim V - 1} v)$$

for all integers $m \geq \dim V - 1$.

Solution. Let $n = \dim V$ and $U = \operatorname{span}(v, Tv, \ldots, T^{n-1}v)$. By Exercise 23, U is invariant under T; since $v \in U$, induction on k gives

$$T^k v \in U \qquad \text{for every integer } k \geq 0.$$

Now fix an integer $m \geq n - 1$. Then

$$U \subseteq \operatorname{span}(v, Tv, \ldots, T^m v) \subseteq U,$$

the first inclusion because $v, Tv, \ldots, T^{n-1}v$ is a sublist of $v, Tv, \ldots, T^m v$, the second because U is a subspace containing each $T^k v$. Hence the two spans are equal.

Problem 5B.11. Suppose V is a two-dimensional vector space, $T \in \mathcal{L}(V)$, and the matrix of T with respect to some basis of V is $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$.
(a) Show that $T^2 - (a + d)T + (ad - bc)I = 0$.
(b) Show that the minimal polynomial of T equals

$$\begin{cases} z - a & \text{if } b = c = 0 \text{ and } a = d, \\ z^2 - (a + d)z + (ad - bc) & \text{otherwise.} \end{cases}$$

Solution. (a) Because $\mathcal{M}$ is linear, injective, and multiplicative (3.43), it suffices to check the identity for $A = \mathcal{M}(T)$, and multiplying out gives

$$A^2 = \begin{pmatrix} a^2 + cb & ac + cd \\ ab + db & bc + d^2 \end{pmatrix} = (a + d)A - (ad - bc)I$$

(compare the four entries). Hence $T^2 - (a+d)T + (ad-bc)I = 0$.
(b) Let p be the minimal polynomial of T and $g(z) = z^2 - (a+d)z + (ad-bc)$, so g is monic of degree 2 with $g(T) = 0$ by (a). Then $\deg p \geq 1$, since the constant polynomial 1 gives $I \neq 0$ (as $\dim V = 2$). Because $\mathcal{M}$ is injective, $T = \lambda I$ holds if and only if $b = c = 0$ and $a = d = \lambda$.
(i) If $b = c = 0$ and $a = d$, then $T = aI$, so the monic degree-1 polynomial $z - a$ annihilates T, and $p(z) = z - a$ by the uniqueness in 5.22.
(ii) Otherwise T is not a scalar multiple of I, so no monic polynomial of degree 1 annihilates T, forcing $\deg p = 2$; since g is monic of degree 2 with $g(T) = 0$, 5.22 gives

$$p(z) = z^2 - (a+d)z + (ad-bc).$$

Problem 5B.12. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, \ldots, x_n) = (x_1, 2x_2, 3x_3, \ldots, nx_n)$. Find the minimal polynomial of T.

Solution. The minimal polynomial of T is

$$(z-1)(z-2)(z-3)\cdots(z-n) = \prod_{j=1}^{n}(z-j).$$

For the standard basis $e_1, \ldots, e_n$ of $\mathbf{F}^n$ the definition of T reads $Te_j = j\,e_j$, so $T^k e_j = j^k e_j$ and hence $q(T)e_j = q(j)e_j$ for every $q \in \mathcal{P}(\mathbf{F})$. As $e_1, \ldots, e_n$ is a basis,

$$q(T) = 0 \iff q(j) = 0 \text{ for each } j = 1, \ldots, n.$$

Thus $g(z) = (z-1)\cdots(z-n)$, monic of degree n, satisfies $g(T) = 0$. The minimal polynomial p has the n numbers $1, \ldots, n$ among its zeros, which are distinct in $\mathbf{F}$ ($= \mathbb{R}$ or $\mathbb{C}$), so $\deg p \geq n$ by 4.8. Hence $\deg p = n$, and $p = g$ by the uniqueness in 5.22.

Problem 5B.13. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $p \in \mathcal{P}(\mathbf{F})$. Prove that there exists a unique $r \in \mathcal{P}(\mathbf{F})$ such that $p(T) = r(T)$ and $\deg r$ is less than the degree of the minimal polynomial of T.

Solution. Take r to be the remainder on dividing p by the minimal polynomial q of T (which exists by 5.22), and set $m = \deg q$.
Existence: by 4.9 there are $s, r \in \mathcal{P}(\mathbf{F})$ with $p = qs + r$ and $\deg r < m$, so applying this to T and using $q(T) = 0$,

$$p(T) = q(T)s(T) + r(T) = r(T).$$

Uniqueness: if r_1, r_2 both qualify, then $u = r_1 - r_2$ satisfies $u(T) = 0$ and $\deg u < m$. By 5.29, $u = qs$ for some $s \in \mathcal{P}(\mathbf{F})$; were $s \neq 0$ we would have $\deg u = m + \deg s \geq m$, a contradiction. Hence $s = 0$, so $u = 0$ and $r_1 = r_2$.

Problem 5B.14. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has minimal polynomial $4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$. Find the minimal polynomial of T^{-1}.

Solution. The minimal polynomial of T^{-1} is the reversal of $p(z) = 4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$, made monic:

$$\tfrac{1}{4} + \tfrac{1}{2}z - \tfrac{7}{4}z^2 - \tfrac{3}{2}z^3 + \tfrac{5}{4}z^4 + z^5.$$

(Here T is invertible by 5.32, the constant term 4 of p being nonzero.)
Reversal lemma: if S is invertible, $u(z) = b_0 + \cdots + b_k z^k$ satisfies $u(S) = 0$, and $\hat{u}(z) = b_k + b_{k-1}z +$

$\cdots + b_0 z^k$, then

$$\begin{aligned}\hat{u}(S^{-1}) &= \sum_{j=0}^{k} b_{k-j} S^{-j} = S^{-k} \sum_{j=0}^{k} b_{k-j} S^{k-j} \\ &= S^{-k} \sum_{i=0}^{k} b_i S^i = S^{-k} u(S) = 0.\end{aligned}$$

Applied with $S = T$ and $u = p$, it gives $\hat{p}(T^{-1}) = 0$ for $\hat{p}(z) = 1 + 2z - 7z^2 - 6z^3 + 5z^4 + 4z^5$; dividing by 4, the monic degree-5 polynomial $g = \frac{1}{4}\hat{p}$ displayed above satisfies $g(T^{-1}) = 0$, so the minimal polynomial $\tilde{q}$ of T^{-1} has degree $k \le 5$.
Conversely, T^{-1} is invertible, so the constant term of $\tilde{q}$ is nonzero by 5.32; hence $\hat{\tilde{q}}$ has degree exactly k, and the lemma with $S = T^{-1}$ gives $\hat{\tilde{q}}(T) = 0$. By 5.29 this nonzero polynomial is a multiple of p, so $k \ge \deg p = 5$.
Thus $k = 5$, and g is monic of degree 5 with $g(T^{-1}) = 0$, so $g = \tilde{q}$ by the uniqueness in 5.22.

Problem 5B.15. Suppose V is a finite-dimensional complex vector space with $\dim V > 0$ and $T \in \mathcal{L}(V)$. Define $f\colon \mathbf{C} \to \mathbf{R}$ by

$$f(\lambda) = \dim \operatorname{range}(T - \lambda I).$$

Prove that f is not continuous.

Solution. The function f fails to be continuous at any eigenvalue λ_0 of T, one of which exists by 5.19 because V is a nonzero finite-dimensional complex vector space.
Write $n = \dim V$. By 3.21,

$$f(\lambda) = n - \dim \operatorname{null}(T - \lambda I),$$

so $f(\lambda) = n$ when λ is not an eigenvalue (then $T - \lambda I$ is injective) while $f(\lambda_0) \le n - 1$. By 5.12 the set of eigenvalues of T is finite, so for each $k \ge 1$ the infinite punctured disk $\{\lambda : 0 < |\lambda - \lambda_0| < 1/k\}$ contains some λ_k that is not an eigenvalue. Then $\lambda_k \to \lambda_0$ and

$$\lim_{k\to\infty} f(\lambda_k) = n > n - 1 \ge f(\lambda_0),$$

so f is not continuous.

Problem 5B.16. Suppose $a_0, \dots, a_{n-1} \in \mathbf{F}$. Let T be the operator on $\mathbf{F}^n$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix}.$$

Here all entries of the matrix are 0 except for all 1's on the line under the diagonal and the entries in the last column (some of which might also be 0). Show that the minimal polynomial of T is the polynomial

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n.$$

The matrix above is called the companion matrix of the polynomial above. This exercise shows that every monic polynomial is the minimal polynomial of some operator. Hence a formula or an algorithm that could produce exact eigenvalues for each operator on each $\mathbf{F}^n$ could then produce exact zeros for each polynomial [by 5.27(a)]. Thus there is no such formula or algorithm. However, efficient numerical methods exist for obtaining very good approximations for the eigenvalues of an operator.

Solution. Write $p(z) = a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} + z^n$. Reading the columns of the displayed matrix against the standard basis $e_1, \ldots, e_n$,

$$Te_j = e_{j+1} \quad (1 \le j \le n-1), \qquad Te_n = -a_0 e_1 - \cdots - a_{n-1}e_n,$$

so induction gives $T^k e_1 = e_{k+1}$ for $k = 0, 1, \ldots, n-1$.
Hence, rewriting each e_{k+1} as $T^k e_1$ in the formula for Te_n,

$$T^n e_1 = Te_n = -\big(a_0 e_1 + a_1 Te_1 + \cdots + a_{n-1}T^{n-1}e_1\big),$$

that is, $p(T)e_1 = 0$. Since $e_j = T^{j-1}e_1$ and powers of T commute with $p(T)$,

$$p(T)e_j = T^{j-1}\big(p(T)e_1\big) = 0 \qquad (j = 1, \ldots, n),$$

so $p(T) = 0$.
No monic q of degree $m < n$ annihilates T: writing $q(z) = b_0 + \cdots + b_{m-1}z^{m-1} + z^m$ and using $T^k e_1 = e_{k+1}$ for $k \le m \le n-1$,

$$q(T)e_1 = b_0 e_1 + b_1 e_2 + \cdots + b_{m-1}e_m + e_{m+1} \ne 0,$$

since the coefficient of e_{m+1} in this combination of the linearly independent list $e_1, \ldots, e_{m+1}$ is 1.
So the minimal polynomial has degree n, and being monic of degree n with $p(T) = 0$, it equals p by the uniqueness in 5.22.

Problem 5B.17. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Suppose $\lambda \in \mathbf{F}$. Show that the minimal polynomial of $T - \lambda I$ is the polynomial q defined by $q(z) = p(z + \lambda)$.

Solution. Everything follows from one shift lemma: for $A \in \mathcal{L}(V)$, $\mu \in \mathbf{F}$, and $r \in \mathcal{P}(\mathbf{F})$ monic of degree m, the polynomial $r_\mu(z) = r(z + \mu)$ is again monic of degree m, and $r_\mu(A) = r(A + \mu I)$.
Indeed, write $r(z) = c_0 + \cdots + c_m z^m$ with $c_m = 1$. Expanding $r_\mu = \sum_k c_k (z + \mu)^k$ by the binomial theorem, the only degree-m contribution is $c_m z^m = z^m$, which gives the first assertion. For the second, $r_\mu = \sum_k c_k s^k$ where $s(z) = z + \mu$, and since $r \mapsto r(A)$ is linear (5.14) and multiplicative (5.17(a)), so that $(s^k)(A) = (A + \mu I)^k$,

$$r_\mu(A) = \sum_{k=0}^{m} c_k (A + \mu I)^k = r(A + \mu I).$$

Now put $m = \deg p$ and $S = T - \lambda I$. Since $q = p_\lambda$, the lemma makes q monic of degree m with

$$q(S) = p(S + \lambda I) = p(T) = 0.$$

Hence the minimal polynomial $\tilde{q}$ of S has $m' = \deg \tilde{q} \le m$. Conversely, $\tilde{q}_{-\lambda}$ is monic of degree m' and

$$\tilde{q}_{-\lambda}(T) = \tilde{q}(T - \lambda I) = \tilde{q}(S) = 0,$$

so minimality of $\deg p$ (5.22) gives $m \le m'$. Thus $m' = m$, and q and $\tilde{q}$ are monic of the same degree m annihilating S, so $q = \tilde{q}$ by the uniqueness in 5.22.

Problem 5B.18. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of

T. Suppose $\lambda \in \mathbf{F} \setminus \{0\}$. Show that the minimal polynomial of λT is the polynomial q defined by

$$q(z) = \lambda^{\deg p}\, p\Big(\frac{z}{\lambda}\Big).$$

Solution. As in Exercise 17, one scaling lemma does all the work: for $A \in \mathcal{L}(V)$, $\mu \in \mathbf{F} \setminus \{0\}$, and $r \in \mathcal{P}(\mathbf{F})$ monic of degree m, the polynomial $r^{[\mu]}(z) = \mu^m r(z/\mu)$ is monic of degree m and $r^{[\mu]}(\mu A) = \mu^m r(A)$.
Indeed, with $r(z) = c_0 + \cdots + c_m z^m$ and $c_m = 1$,

$$r^{[\mu]}(z) = \mu^m \sum_{k=0}^{m} c_k \mu^{-k} z^k = \sum_{k=0}^{m} c_k \mu^{m-k} z^k,$$

whose coefficient of z^m is $c_m = 1$; and since $(\mu A)^k = \mu^k A^k$,

$$r^{[\mu]}(\mu A) = \sum_{k=0}^{m} c_k \mu^{m-k} \mu^k A^k = \mu^m r(A).$$

Let $m = \deg p$, so the exercise's q is $p^{[\lambda]}$, monic of degree m with $q(\lambda T) = \lambda^m p(T) = 0$. Hence the minimal polynomial $\tilde{q}$ of λT has $m' = \deg \tilde{q} \le m$ by 5.22. Conversely, applying the lemma with $\mu = 1/\lambda$ and $A = \lambda T$, the polynomial $\tilde{q}^{[1/\lambda]}$ is monic of degree m' and, since $\frac{1}{\lambda}(\lambda T) = T$,

$$\tilde{q}^{[1/\lambda]}(T) = \lambda^{-m'}\, \tilde{q}(\lambda T) = 0,$$

so minimality of $\deg p$ gives $m \le m'$. Thus $m' = m$, and $q = \tilde{q}$ by the uniqueness in 5.22.

Problem 5B.19. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\mathcal{E}$ be the subspace of $\mathcal{L}(V)$ defined by

$$\mathcal{E} = \{q(T) : q \in \mathcal{P}(\mathbf{F})\}.$$

Prove that $\dim \mathcal{E}$ equals the degree of the minimal polynomial of T.

Solution. The list $I, T, \ldots, T^{m-1}$ is a basis of $\mathcal{E}$, where p is the minimal polynomial of T and $m = \deg p$; hence $\dim \mathcal{E} = m$.
Each T^k lies in $\mathcal{E}$ (take $q(z) = z^k$). The list spans: given $q \in \mathcal{P}(\mathbf{F})$, divide by p using 4.9 to get $q = ps + r$ with $\deg r < m$, so by 5.17(a) and $p(T) = 0$,

$$q(T) = p(T)s(T) + r(T) = r(T) = b_0 I + b_1 T + \cdots + b_{m-1} T^{m-1},$$

writing $r(z) = b_0 + \cdots + b_{m-1} z^{m-1}$.
The list is linearly independent: if $c_0 I + \cdots + c_{m-1} T^{m-1} = 0$ with some $c_k \ne 0$, let j be the largest index with $c_j \ne 0$; dividing by c_j gives

$$r(z) = \frac{c_0}{c_j} + \cdots + \frac{c_{j-1}}{c_j} z^{j-1} + z^j,$$

a monic polynomial of degree $j \le m - 1$ with $r(T) = 0$, contradicting the minimality of $\deg p$ in 5.22.

Problem 5B.20. Suppose $T \in \mathcal{L}(\mathbf{F}^4)$ is such that the eigenvalues of T are $3, 5, 8$. Prove that

$$(T - 3I)^2 (T - 5I)^2 (T - 8I)^2 = 0.$$

Solution. It suffices to show that $r(z) = (z-3)^2(z-5)^2(z-8)^2$ is a polynomial multiple of the minimal polynomial p of T, since then $r(T) = 0$ by 5.29, and $r(T) = (T - 3I)^2(T - 5I)^2(T - 8I)^2$ by 5.17(a).

By 5.22, $\deg p \le \dim \mathbf{F}^4 = 4$, and by 5.27(a) the zeros of p are exactly the eigenvalues $3, 5, 8$. Factoring these out one at a time by 4.6 (legitimate because each successive monic quotient is nonzero and vanishes at the next of the three numbers: writing $p(z) = (z-3)q(z)$, the equation $0 = p(5) = (5-3)q(5)$ forces $q(5) = 0$, and similarly at 8) gives

$$p(z) = (z-3)(z-5)(z-8)\, t(z)$$

with t monic and $\deg t = \deg p - 3 \le 1$.
(i) $\deg t = 0$: then $p(z) = (z-3)(z-5)(z-8)$ and $r(z) = p(z)(z-3)(z-5)(z-8)$.
(ii) $\deg t = 1$: then $t(z) = z - c$ with c a zero of p, so $c \in \{3, 5, 8\}$ by 5.27(a). Writing d, e for the other two members of $\{3, 5, 8\}$,

$$p(z) = (z-c)^2(z-d)(z-e), \qquad r(z) = p(z)\,(z-d)(z-e).$$

Problem 5B.21. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of T has degree at most $1 + \dim \operatorname{range} T$.
If $\dim \operatorname{range} T < \dim V - 1$, then this exercise gives a better upper bound than 5.22 for the degree of the minimal polynomial of T.

Solution. Put $U = \operatorname{range} T$, which is invariant under T (if $u \in U$ then $Tu \in \operatorname{range} T = U$), and let q be the minimal polynomial of $T|_U \in \mathcal{L}(U)$, so that q is monic with $\deg q \le \dim U$ by 5.22.
Because U is invariant, induction on k gives $T^k u = (T|_U)^k u$ for all $u \in U$ and $k \ge 0$, hence

$$q(T)u = q(T|_U)u = 0 \qquad \text{for all } u \in U.$$

Now $s(z) = z\, q(z)$ is monic with $\deg s = 1 + \deg q \le 1 + \dim U$, and $s(T) = q(T)\, T$ by 5.17(a), so for every $v \in V$,

$$s(T)v = q(T)(Tv) = 0,$$

the last step by the previous display, since $Tv \in U$. Thus s is a monic polynomial annihilating T, so by the minimality in 5.22 the minimal polynomial p of T satisfies

$$\deg p \le \deg s \le 1 + \dim \operatorname{range} T.$$

Problem 5B.22. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if $I \in \operatorname{span}(T, T^2, \dots, T^{\dim V})$.

Solution. Let $n = \dim V$.
Suppose T is invertible, and let $p(z) = c_0 + c_1 z + \cdots + c_{m-1} z^{m-1} + z^m$ be its minimal polynomial, so $m \le n$ by 5.22 and $c_0 \ne 0$ by 5.32. Dividing $p(T) = 0$ by c_0 and rearranging,

$$I = -\frac{c_1}{c_0} T - \cdots - \frac{c_{m-1}}{c_0} T^{m-1} - \frac{1}{c_0} T^m,$$

which lies in $\operatorname{span}(T, T^2, \dots, T^n)$ because $m \le n$. (If $m = 0$ then $p = 1$ forces $I = 0$ and $V = \{0\}$, where I lies in every span.)
Conversely, suppose $I = a_1 T + a_2 T^2 + \cdots + a_n T^n$ and set $S = a_1 I + a_2 T + \cdots + a_n T^{n-1}$. Powers of T commute, so factoring T out on either side gives

$$TS = ST = a_1 T + a_2 T^2 + \cdots + a_n T^n = I,$$

so S is an inverse of T.

Problem 5B.23. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Prove that if $v \in V$, then $\operatorname{span}(v, Tv, \ldots, T^{n-1}v)$ is invariant under T.

Solution. Put $U = \operatorname{span}(v, Tv, \ldots, T^{n-1}v)$. Since $T(T^k v) = T^{k+1}v \in U$ for $k \leq n-2$, the only point to prove is $T^n v \in U$.
Let m be the smallest nonnegative integer with $T^m v \in W$, where

$$W = \operatorname{span}(v, Tv, \ldots, T^{m-1}v).$$

Such an $m \leq n$ exists: the list $v, Tv, \ldots, T^n v$ has length $n+1 > \dim V$, so it is linearly dependent by 2.22, and then 2.19 puts one of its terms in the span of the preceding terms.
Now W is invariant under T: for $j \leq m-2$ we have $T(T^j v) = T^{j+1}v \in W$, while $T(T^{m-1}v) = T^m v \in W$ by the choice of m. Since $v \in W$ (when $m = 0$, $W = \{0\}$ and $v = 0$), induction gives $T^k v \in W$ for every $k \geq 0$.
(i) $m \leq n-1$: then $W \subseteq U$ because the spanning list of W is a sublist of that of U, and $U \subseteq W$ by the previous display, so $U = W$ is invariant under T.
(ii) $m = n$: minimality of m together with 2.19 makes $v, Tv, \ldots, T^{n-1}v$ linearly independent, hence a basis of V by 2.38 (its length is $n = \dim V$), so $U = V$ is invariant under T.

Problem 5B.24. Suppose V is a finite-dimensional complex vector space. Suppose $T \in \mathcal{L}(V)$ is such that 5 and 6 are eigenvalues of T and that T has no other eigenvalues. Prove that $(T-5I)^{\dim V - 1}(T-6I)^{\dim V-1} = 0$.

Solution. Let $n = \dim V$ and let p be the minimal polynomial of T, so $\deg p \leq n$ by 5.22. Since $\mathbf{F} = \mathbf{C}$, part (b) of 5.27 factors p into linear factors, and by part (a) its zeros are exactly the eigenvalues 5 and 6; hence

$$p(z) = (z-5)^j (z-6)^k, \qquad j, k \geq 1, \quad j + k = \deg p \leq n.$$

Therefore $j \leq n-1$ and $k \leq n-1$, so the exponents below are nonnegative and

$$(z-5)^{n-1}(z-6)^{n-1} = p(z)\,(z-5)^{n-1-j}(z-6)^{n-1-k}.$$

Being a polynomial multiple of p, this polynomial evaluates to 0 at T by 5.29, that is, $(T-5I)^{n-1}(T-6I)^{n-1} = 0$.

Problem 5B.25. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T.
(a) Prove that the minimal polynomial of T is a polynomial multiple of the minimal polynomial of the quotient operator T/U.
(b) Prove that

$$(\text{minimal polynomial of } T|_U) \times (\text{minimal polynomial of } T/U)$$

is a polynomial multiple of the minimal polynomial of T.
[The quotient operator T/U was defined in Exercise 38 in Section 5A.]

Solution. Two computations drive both parts. Since $(T/U)(v+U) = Tv + U$ (Exercise 38 in Section 5A), induction gives $(T/U)^k(v+U) = T^k v + U$, so by linearity of the quotient map (3.104),

$$q(T/U)(v+U) = q(T)v + U \qquad (q \in \mathcal{P}(\mathbf{F}),\ v \in V). \quad (*)$$

Likewise $T^k u = (T|_U)^k u$ for $u \in U$ (induction, using invariance of U), so

$$q(T)u = q(T|_U)u \qquad (q \in \mathcal{P}(\mathbf{F}),\ u \in U). \quad (**)$$

Both $T|_U$ and T/U have minimal polynomials by 5.22, since U and V/U are finite-dimensional (3.105).

(a) Let p be the minimal polynomial of T. By $(*)$, for every $v \in V$,

$$p(T/U)(v+U) = p(T)v + U = 0 + U,$$

so $p(T/U) = 0$ and 5.29, applied to T/U, makes p a polynomial multiple of the minimal polynomial of T/U.
(b) Let r and s be the minimal polynomials of $T|_U$ and T/U. By 5.29 it suffices to prove $(rs)(T) = 0$. Fix $v \in V$. Applying $(*)$ with $q = s$ and using $s(T/U) = 0$ gives $s(T)v \in U$, so $(**)$ yields

$$(rs)(T)v = r(T)\big(s(T)v\big) = r(T|_U)\big(s(T)v\big) = 0,$$

the first equality by 5.17 and the last because $r(T|_U) = 0$.

Problem 5B.26. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T. Prove that the set of eigenvalues of T equals the union of the set of eigenvalues of $T|_U$ and the set of eigenvalues of T/U.

Solution. Let p, r, s be the minimal polynomials of T, $T|_U$, T/U (all exist by 5.22, as V, U, V/U are finite-dimensional), and write $Z(q)$ for the set of zeros of q in $\mathbf{F}$. By 5.27(a) these zero sets are the respective sets of eigenvalues, so it suffices to prove $Z(p) = Z(r) \cup Z(s)$.
By 5.31 and by Exercise 25(a), p is a polynomial multiple of r and of s; hence every zero of r or of s is a zero of p, giving $Z(r) \cup Z(s) \subseteq Z(p)$.
Conversely, Exercise 25(b) gives $rs = pk$ for some $k \in \mathcal{P}(\mathbf{F})$, so if $\lambda \in Z(p)$ then

$$r(\lambda)s(\lambda) = p(\lambda)k(\lambda) = 0,$$

and $\mathbf{F}$ has no zero divisors, so $r(\lambda) = 0$ or $s(\lambda) = 0$. Hence $Z(p) \subseteq Z(r) \cup Z(s)$.

Problem 5B.27. Suppose $\mathbf{F} = \mathbf{R}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T_{\mathbf{C}}$ equals the minimal polynomial of T.
[The complexification $T_{\mathbf{C}}$ was defined in Exercise 33 of Section 3B.]

Solution. Let $p \in \mathcal{P}(\mathbf{R})$ be the minimal polynomial of T. A basis $v_1, \ldots, v_n$ of V over $\mathbf{R}$ is also a basis of $V_{\mathbf{C}}$ over $\mathbf{C}$ (Check!, using $u + iv = \sum_j (a_j + ib_j)v_j$ and coordinatewise equality), so $V_{\mathbf{C}}$ is finite-dimensional and $T_{\mathbf{C}}$ has a minimal polynomial $q \in \mathcal{P}(\mathbf{C})$ by 5.22. We show p and q are multiples of each other.
For $g \in \mathcal{P}(\mathbf{R})$,

$$g(T_{\mathbf{C}})(u+iv) = g(T)u + i\,g(T)v \qquad (u, v \in V), \quad (\dagger)$$

since $(T_{\mathbf{C}})^k(u+iv) = T^k u + iT^k v$ by induction and $a(x+iy) = ax + iay$ for real a.
Taking $g = p$ in $(\dagger)$ gives $p(T_{\mathbf{C}}) = 0$, so p is a polynomial multiple of q by 5.29.
Conversely, write $q = q_1 + iq_2$ with $q_1, q_2 \in \mathcal{P}(\mathbf{R})$ (real and imaginary parts of the coefficients). For $u \in V$, $(\dagger)$ applied to q_1 and q_2 gives

$$0 = q(T_{\mathbf{C}})(u + i0) = q_1(T)u + i\,q_2(T)u,$$

and equality in $V_{\mathbf{C}} = V \times V$ is coordinatewise, so $q_1(T) = q_2(T) = 0$. By 5.29 each q_j is a polynomial multiple of p, say $q_j = pg_j$, whence $q = p(g_1 + ig_2)$.
Thus $q = ph$ and $p = qg$, so $p = phg$; as $p \neq 0$, h is a nonzero constant, and comparing leading coefficients of the monic p and q gives $h = 1$. Hence $q = p$.

Problem 5B.28. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T' \in \mathcal{L}(V')$ equals the minimal polynomial of T.

[The dual map T' was defined in Section 3F.]

Solution. The two minimal polynomials agree because T and T' are annihilated by exactly the same polynomials. (Both exist: $\dim V' = \dim V < \infty$ by 3.111, so 5.22 applies to each.)
First, $(T^k)' = (T')^k$ for every integer $k \geq 0$: for $k = 0$ this is $I'(\varphi) = \varphi \circ I = \varphi$, and inductively 3.120(c) gives

$$\begin{aligned}(T^{k+1})' &= (T^k T)' = T'\,(T^k)' \\ &= T'(T')^k = (T')^{k+1}.\end{aligned}$$

Since $S \mapsto S'$ is linear by 3.120(a) and 3.120(b), applying this termwise to $g(z) = a_0 + a_1 z + \cdots + a_d z^d$ yields $\big(g(T)\big)' = g(T')$ for every $g \in \mathcal{P}(\mathbf{F})$.
Next, $S = 0$ if and only if $S' = 0$, because $\dim \operatorname{range} S' = \dim \operatorname{range} S$ by 3.130(a) and an operator vanishes exactly when its range has dimension 0. Combining the two facts,

$$g(T) = 0 \iff \big(g(T)\big)' = 0 \iff g(T') = 0.$$

So if p and $\tilde{p}$ are the minimal polynomials of T and T', then $p(T') = 0$ gives $\deg \tilde{p} \leq \deg p$ and $\tilde{p}(T) = 0$ gives $\deg p \leq \deg \tilde{p}$. Thus $\tilde{p}$ is monic of degree $\deg p$ with $\tilde{p}(T) = 0$, so $\tilde{p} = p$ by the uniqueness in 5.22.

Problem 5B.29. Show that every operator on a finite-dimensional vector space of dimension at least two has an invariant subspace of dimension two.
[Exercise 6 in Section 5C will give an improvement of this result when $\mathbf{F} = \mathbf{C}$.]

Solution. Take $W = \operatorname{span}(w, Tw)$, where $w = h(T)v$ for the vector v and the factorization $q = gh$ produced below.
(i) If $T = \lambda I$, every subspace of V is invariant under T, and $\dim V \geq 2$ supplies a two-dimensional one.
(ii) Otherwise some $v \in V$ has v, Tv linearly independent. For if not, then every nonzero $u \in V$ satisfies $Tu = \lambda_u u$ for some $\lambda_u \in \mathbf{F}$ (in a dependence $au + bTu = 0$ we must have $b \neq 0$, else $u = 0$), and all these scalars agree: if $w = tu$ with $t \neq 0$, then $\lambda_w w = t\lambda_u u = \lambda_u w$, while if u, w is linearly independent, then comparing coefficients in $\lambda_{u+w} u + \lambda_{u+w} w = \lambda_u u + \lambda_w w$ gives $\lambda_u = \lambda_{u+w} = \lambda_w$. Thus $T = \lambda I$, a contradiction.
Fix such a v and let q be the monic polynomial of smallest degree m with $q(T)v = 0$, which exists by Exercise 7(a) in Section 5C. Minimality gives the consequence, used twice below, that $r(T)v = 0$ with $\deg r < m$ forces $r = 0$ (otherwise divide r by its leading coefficient). Also $m \geq 2$, since $m = 0$ would give $v = Iv = 0$ and $m = 1$ would give $Tv = \lambda v$, contradicting the independence of v, Tv.
Any monic q with $\deg q = m \geq 2$ factors as $q = gh$ with g, h monic, $\deg g = 2$ and $\deg h = m - 2$: over $\mathbf{C}$ let g be the product of the first two linear factors supplied by 4.13, and over $\mathbf{R}$ let g be a quadratic factor $x^2 + b_1 x + c_1$ supplied by 4.16 if one occurs there and otherwise the product of two of the m linear factors. Write $g(z) = z^2 + bz + c$.
Then $w = h(T)v \neq 0$, since h is monic of degree $m - 2 < m$, hence nonzero. By 5.17,

$$g(T)w = (gh)(T)v = q(T)v = 0, \qquad \text{so} \qquad T^2 w = -b\,Tw - c\,w,$$

whence $W = \operatorname{span}(w, Tw)$ is invariant under T, because

$$T(\alpha w + \beta\,Tw) = (-\beta c)w + (\alpha - \beta b)\,Tw \in W.$$

Finally $\dim W = 2$: otherwise $W = \operatorname{span}(w)$ with $w \neq 0$, so $Tw = \lambda w$ for some $\lambda \in \mathbf{F}$, and then $r(z) = (z - \lambda)h(z)$ is monic of degree $m - 1$ with $r(T)v = (T - \lambda I)w = 0$ by 5.17, contradicting the minimality of m.

5.3 Exercises 5C

Problem 5C.1. Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and T^2 has an upper-triangular matrix with respect to some basis of V, then T has an upper-triangular matrix with respect to some basis of V.

Solution. False: take $\mathbf{F} = \mathbf{R}$, $V = \mathbf{R}^2$, and $T(x, y) = (-y, x)$, rotation by 90 degrees.
Then $T^2(x, y) = T(-y, x) = (-x, -y)$, so $T^2 = -I$, whose matrix with respect to every basis of $\mathbf{R}^2$ is

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is upper triangular.
But T has no eigenvalue: if $T(x, y) = \lambda(x, y)$, then $-y = \lambda x$ and $x = \lambda y$, so $x = -\lambda^2 x$, hence $(1 + \lambda^2)x = 0$ and $x = 0$, and then $y = -\lambda x = 0$ too. Were the matrix of T with respect to some basis upper triangular, its diagonal entries would be eigenvalues of T by 5.41, a contradiction.

Problem 5C.2. Suppose A and B are upper-triangular matrices of the same size, with $\alpha_1, \ldots, \alpha_n$ on the diagonal of A and $\beta_1, \ldots, \beta_n$ on the diagonal of B.

- (a) Show that $A + B$ is an upper-triangular matrix with $\alpha_1 + \beta_1, \ldots, \alpha_n + \beta_n$ on the diagonal.
- (b) Show that AB is an upper-triangular matrix with $\alpha_1\beta_1, \ldots, \alpha_n\beta_n$ on the diagonal.

[The results in this exercise are used in the proof of 5.81.]

Solution. Both claims are entrywise consequences of $A_{j,k} = B_{j,k} = 0$ whenever $j > k$ (5.38), with $A_{k,k} = \alpha_k$ and $B_{k,k} = \beta_k$.
(a) $(A + B)_{j,k} = A_{j,k} + B_{j,k}$, which is 0 when $j > k$ and equals $\alpha_k + \beta_k$ when $j = k$.
(b) By 3.41, $(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r}B_{r,k}$, and a nonzero term forces $A_{j,r} \neq 0$ and $B_{r,k} \neq 0$, hence

$$j \leq r \leq k.$$

If $j > k$, no such r exists, so $(AB)_{j,k} = 0$; if $j = k$, the only such r is k, so $(AB)_{k,k} = A_{k,k}B_{k,k} = \alpha_k\beta_k$.

Problem 5C.3. Suppose $T \in \mathcal{L}(V)$ is invertible and $v_1, \ldots, v_n$ is a basis of V with respect to which the matrix of T is upper triangular, with $\lambda_1, \ldots, \lambda_n$ on the diagonal. Show that the matrix of T^{-1} is also upper triangular with respect to the basis $v_1, \ldots, v_n$, with

$$\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}$$

on the diagonal.

Solution. Each $\lambda_k \neq 0$, because T is invertible and hence injective, so 0 is not an eigenvalue of T, while by 5.41 the eigenvalues of T are exactly $\lambda_1, \ldots, \lambda_n$.
Fix k and set $U_k = \operatorname{span}(v_1, \ldots, v_k)$, which is invariant under T by 5.39. Injectivity of T makes the operator $T|_{U_k}$ on the finite-dimensional space U_k injective, hence surjective by 3.65, so $T(U_k) = U_k$. Applying T^{-1} to $v_k \in U_k = T(U_k)$ therefore gives

$$T^{-1}v_k \in \operatorname{span}(v_1, \ldots, v_k) \qquad \text{for each } k,$$

so the matrix of T^{-1} with respect to $v_1, \ldots, v_n$ is upper triangular by the implication (c) $\Longrightarrow$ (a) of 5.39.
Let $\mu_1, \ldots, \mu_n$ be its diagonal entries. By 3.43 (all three bases equal to $v_1, \ldots, v_n$),

$$\mathcal{M}(T)\,\mathcal{M}(T^{-1}) = \mathcal{M}(TT^{-1}) = \mathcal{M}(I),$$

whose diagonal entries all equal 1, while Exercise 2(b) says that diagonal is $\lambda_1\mu_1, \ldots, \lambda_n\mu_n$. Hence

$\lambda_k \mu_k = 1$, that is, $\mu_k = 1/\lambda_k$.

Problem 5C.4. Give an example of an operator whose matrix with respect to some basis contains only 0's on the diagonal, but the operator is invertible.
[This exercise and the exercise below show that 5.41 fails without the hypothesis that an upper-triangular matrix is under consideration.]

Solution. Take $T \in \mathcal{L}(\mathbf{F}^2)$ with $T(x, y) = (y, x)$. With respect to the standard basis e_1, e_2,

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

whose diagonal entries are both 0, yet $T\big(T(x, y)\big) = T(y, x) = (x, y)$, so $T^2 = I$ and T is invertible with $T^{-1} = T$.

Problem 5C.5. Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.

Solution. Take $T \in \mathcal{L}(\mathbf{F}^2)$ with $T(x, y) = (x + y,\ x + y)$. With respect to the standard basis e_1, e_2,

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

whose diagonal entries both equal 1, yet $T(1, -1) = (0, 0)$ with $(1, -1) \neq 0$, so T is not injective and hence not invertible.

Problem 5C.6. Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that if $k \in \{1, \ldots, \dim V\}$, then V has a k-dimensional subspace invariant under T.

Solution. Take $U = \operatorname{span}(v_1, \ldots, v_k)$, where $v_1, \ldots, v_n$ is a basis of V with respect to which T has an upper-triangular matrix, supplied by 5.47 because V is a finite-dimensional complex vector space. Then U is invariant under T by the implication (a) $\Longrightarrow$ (b) of 5.39, and $v_1, \ldots, v_k$ is a linearly independent spanning list of U, being a sublist of a basis; hence $\dim U = k$.

Problem 5C.7. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$.

- (a) Prove that there exists a unique monic polynomial p_v of smallest degree such that $p_v(T)v = 0$.
- (b) Prove that the minimal polynomial of T is a polynomial multiple of p_v.

Solution. (a) Take p_v to be a member of smallest degree m of

$$S = \{\, q \in \mathcal{P}(\mathbf{F}) : q \text{ is monic and } q(T)v = 0 \,\},$$

which exists because S contains the minimal polynomial p of T (as $p(T) = 0$ by 5.22, V being finite-dimensional), so $\{\deg q : q \in S\}$ is a nonempty set of nonnegative integers.
For uniqueness, suppose $q \in S$ also has degree m. The leading terms of the two monic polynomials cancel, so $p_v - q = 0$ or $\deg(p_v - q) < m$, and $(p_v - q)(T)v = 0$ by the linearity of $r \mapsto r(T)$ (5.14). Were $p_v - q \neq 0$ with leading coefficient c, the polynomial $c^{-1}(p_v - q)$ would lie in S with degree less than m, contradicting minimality. Hence $q = p_v$.
(b) Since $p_v \neq 0$, the division algorithm 4.9 gives $p = s\,p_v + r$ with $\deg r < \deg p_v$. Applying both sides to v, using 5.14 and 5.17(a) together with $p(T) = 0$ and $p_v(T)v = 0$,

$$0 = p(T)v = s(T)\big(p_v(T)v\big) + r(T)v = r(T)v.$$

If $r \neq 0$ with leading coefficient c, then $c^{-1}r$ is monic of degree $\deg r < \deg p_v$ with $(c^{-1}r)(T)v = 0$, contradicting (a). So $r = 0$ and $p = s\,p_v$.

Problem 5C.8. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and there exists a nonzero vector $v \in V$ such that $T^2v + 2Tv = -2v$.
(a) Prove that if $\mathbb{F} = \mathbb{R}$, then there does not exist a basis of V with respect to which T has an upper-triangular matrix.
(b) Prove that if $\mathbb{F} = \mathbb{C}$ and A is an upper-triangular matrix that equals the matrix of T with respect to some basis of V, then $-1+i$ or $-1-i$ appears on the diagonal of A.

Solution. Both parts run through the nonzero T-invariant subspace $U = \operatorname{span}(v, Tv)$, on which $S = T|_U$ satisfies $S^2 + 2S + 2I_U = 0$.
The hypothesis reads $(T^2 + 2T + 2I)v = 0$ with $v \neq 0$, so $T(Tv) = -2Tv - 2v \in U$ and U is invariant under T; moreover $T^2 + 2T + 2I$ also kills Tv (it commutes with T), hence kills all of U, which is the displayed identity. So if $\mu \in \mathbb{F}$ is an eigenvalue of S, with eigenvector $u \neq 0$, then

$$0 = (S^2 + 2S + 2I_U)u = (\mu^2 + 2\mu + 2)u,$$

so $(\mu + 1)^2 = -1$.
In either part, an upper-triangular matrix for T with diagonal $\lambda_1, \dots, \lambda_n$ gives $(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0$ by 5.40; since U is invariant under T and hence under each $T - \lambda_k I$, restricting to U yields

$$(S - \lambda_1 I_U) \cdots (S - \lambda_n I_U) = 0.$$

(a) $\mathbb{F} = \mathbb{R}$: no real μ satisfies $(\mu + 1)^2 = -1$, so S has no eigenvalue and every factor $S - \lambda_k I_U$ is injective by 5.7. Then the composition is injective, so the display forces $U = \{0\}$, contradicting $v \neq 0$.
(b) $\mathbb{F} = \mathbb{C}$: here $(\mu+1)^2 = -1$ forces $\mu = -1+i$ or $\mu = -1-i$. Since $U \neq \{0\}$, the composition in the display is not injective, so some factor $S - \lambda_k I_U$ is not injective; by 5.7 that λ_k is an eigenvalue of S, hence equals $-1+i$ or $-1-i$, and it lies on the diagonal of A.

Problem 5C.9. Suppose B is a square matrix with complex entries. Prove that there exists an invertible square matrix A with complex entries such that $A^{-1}BA$ is an upper-triangular matrix.

Solution. Take $A = \mathcal{M}\big(I, (u_1, \dots, u_n), (e_1, \dots, e_n)\big)$, the change-of-basis matrix described below.
Say B is n-by-n, let $e_1, \dots, e_n$ be the standard basis of $\mathbb{C}^n$, and let $T \in \mathcal{L}(\mathbb{C}^n)$ be the operator with $Te_k = \sum_{j=1}^n B_{j,k}\, e_j$, so that $\mathcal{M}\big(T, (e_1, \dots, e_n)\big) = B$. Because $\mathbb{C}^n$ is a finite-dimensional complex vector space, 5.47 provides a basis $u_1, \dots, u_n$ of $\mathbb{C}^n$ for which

$$C = \mathcal{M}\big(T, (u_1, \dots, u_n)\big)$$

is upper triangular.
This A is invertible: with $A' = \mathcal{M}\big(I, (e_1, \dots, e_n), (u_1, \dots, u_n)\big)$, the product formula 3.81 gives $AA' = \mathcal{M}(I, (e), (e))$ and $A'A = \mathcal{M}(I, (u), (u))$, both the identity matrix. The change-of-basis formula 3.84 then gives

$$A^{-1}BA = C,$$

which is upper triangular.

Problem 5C.10. Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Show that the following are equivalent.
(a) The matrix of T with respect to $v_1, \dots, v_n$ is lower triangular.
(b) $\operatorname{span}(v_k, \dots, v_n)$ is invariant under T for each $k = 1, \dots, n$.
(c) $Tv_k \in \operatorname{span}(v_k, \dots, v_n)$ for each $k = 1, \dots, n$.
[A square matrix is called lower triangular if all entries above the diagonal are 0.]

Solution. Write $A = \mathcal{M}\big(T, (v_1, \dots, v_n)\big)$, so that $Tv_k = \sum_{i=1}^n A_{i,k} v_i$, and prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b). Lower triangularity says $A_{i,j} = 0$ whenever $i < j$, so

$$\begin{aligned} Tv_j &\in \operatorname{span}(v_j, \dots, v_n) \\ &\subseteq \operatorname{span}(v_k, \dots, v_n) \qquad \text{for } j \geq k. \end{aligned}$$

Thus T maps each vector of the spanning list $v_k, \dots, v_n$ into $\operatorname{span}(v_k, \dots, v_n)$, hence maps that subspace into itself.
(b) $\Rightarrow$ (c). v_k lies in $\operatorname{span}(v_k, \dots, v_n)$, which is invariant under T.
(c) $\Rightarrow$ (a). Writing Tv_k as a combination of $v_k, \dots, v_n$ alone says $A_{i,k} = 0$ for all $i < k$, so every entry of A above the diagonal is 0.

Problem 5C.11. Suppose $\mathbb{F} = \mathbb{C}$ and V is finite-dimensional. Prove that if $T \in \mathcal{L}(V)$, then there exists a basis of V with respect to which T has a lower-triangular matrix.

Solution. Reverse a triangularizing basis: let $v_1, \dots, v_n$ be a basis of V with respect to which T has an upper-triangular matrix, supplied by 5.47 because V is a finite-dimensional complex vector space, and put $w_j = v_{n+1-j}$.
Then $w_1, \dots, w_n$ is again a basis of V, being the same vectors in a different order, and 5.39 gives $Tv_m \in \operatorname{span}(v_1, \dots, v_m)$ for each m. Fix k and set $m = n + 1 - k$, so $w_k = v_m$; since span does not depend on the order of the list,

$$\begin{aligned} \operatorname{span}(w_k, \dots, w_n) &= \operatorname{span}(v_{n+1-k}, \dots, v_1) \\ &= \operatorname{span}(v_1, \dots, v_m). \end{aligned}$$

Hence $Tw_k = Tv_m \in \operatorname{span}(w_k, \dots, w_n)$, which is condition (c) of Exercise 10; so by Exercise 10 the matrix of T with respect to $w_1, \dots, w_n$ is lower triangular.

Problem 5C.12. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V, and U is a subspace of V that is invariant under T.
(a) Prove that $T|_U$ has an upper-triangular matrix with respect to some basis of U.
(b) Prove that the quotient operator T/U has an upper-triangular matrix with respect to some basis of V/U.
[The quotient operator T/U was defined in Exercise 38 in Section 5A.]

Solution. Both parts come from the flag $V_k = \operatorname{span}(v_1, \dots, v_k)$ of a basis $v_1, \dots, v_n$ of V with respect to which T has an upper-triangular matrix: each V_k is invariant under T by 5.39 and has $\dim V_k = k$, with $V_0 = \{0\}$ and $V_n = V$. Part (a) intersects this chain with U; part (b) pushes it into V/U.
(a) Set $U_k = U \cap V_k$, invariant under T because U and V_k both are, with $U_0 = \{0\}$, $U_n = U$, and $U_{k-1} \subseteq U_k$. Since $U_k \cap V_{k-1} = U \cap V_{k-1} = U_{k-1}$ while $U_k + V_{k-1} \subseteq V_k$, the dimension formula 2.43 gives

$$\dim U_k - \dim U_{k-1} + (k-1) = \dim(U_k + V_{k-1}) \leq k,$$

so each step raises the dimension by 0 or 1. Let $m = \dim U$ and let $k_1 < \cdots < k_m$ be the indices at which it rises; counting the rises among $1, \dots, k_j$ gives $\dim U_{k_j} = j$.
Choose $u_j \in U_{k_j}$ with $u_j \notin U_{k_j - 1}$, possible since that inclusion is strict. Then $\operatorname{span}(u_1, \dots, u_j) = U_{k_j}$ for each j, by induction: $U_{k_{j-1}} \subseteq U_{k_j - 1}$ and $u_j \notin U_{k_j-1}$ make $u_1, \dots, u_j$ linearly independent, and a linearly independent list of length j in the j-dimensional space U_{k_j} is a basis of it (2.38). Taking $j = m$ makes $u_1, \dots, u_m$ a basis of U, and each $\operatorname{span}(u_1, \dots, u_j) = U_{k_j}$ is contained in U and invariant under T, hence under $T|_U$; so the implication (b) $\Longrightarrow$ (a) of 5.39, applied to $T|_U$, makes the matrix of $T|_U$ with respect to $u_1, \dots, u_m$ upper triangular.
(b) Let $\pi \colon V \to V/U$ be the quotient map and $W_k = \pi(V_k) = \operatorname{span}\big(\pi(v_1), \dots, \pi(v_k)\big)$. Each W_k is invariant under T/U, since $(T/U)\big(\pi(v_i)\big) = \pi(Tv_i) \in \pi(V_k) = W_k$ for $i \leq k$. As π is surjective, $\pi(v_1), \dots, \pi(v_n)$ spans V/U; reduce it to a basis by the procedure of 2.30, deleting $\pi(v_k)$ whenever it lies in the span of its predecessors, and let $k_1 < \cdots < k_p$ be the surviving indices, $w_j = \pi(v_{k_j})$.

Each deletion leaves the span of every initial segment unchanged, so induction on k gives

$$\operatorname{span}\big(w_j : k_j \leq k\big) = \operatorname{span}\big(\pi(v_1), \dots, \pi(v_k)\big) = W_k.$$

Taking $k = k_j$ gives $\operatorname{span}(w_1, \dots, w_j) = W_{k_j}$, which is invariant under T/U; so 5.39, applied to T/U and the basis $w_1, \dots, w_p$ of V/U, makes that matrix upper triangular.

Problem 5C.13. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose there exists a subspace U of V that is invariant under T such that $T|_U$ has an upper-triangular matrix with respect to some basis of U and also T/U has an upper-triangular matrix with respect to some basis of V/U. Prove that T has an upper-triangular matrix with respect to some basis of V.

Solution. Take the concatenated list $x_1, \dots, x_{m+p}$ equal to $u_1, \dots, u_m, w_1, \dots, w_p$, where $u_1, \dots, u_m$ is a basis of U triangularizing $T|_U$, the list $\widetilde{w}_1, \dots, \widetilde{w}_p$ is a basis of V/U triangularizing T/U, and $\pi(w_j) = \widetilde{w}_j$ for the quotient map $\pi\colon V \to V/U$, which is surjective with $\operatorname{null} \pi = U$ and $(T/U) \circ \pi = \pi \circ T$.
This list is a basis of V. It spans V: given $v \in V$, write $\pi(v) = c_1\widetilde{w}_1 + \cdots + c_p\widetilde{w}_p$; then $\pi(v - c_1 w_1 - \cdots - c_p w_p) = 0$, so that vector lies in $\operatorname{null} \pi = U = \operatorname{span}(u_1, \dots, u_m)$. It is linearly independent: applying π to $a_1u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_p w_p = 0$ kills the u_i and leaves $b_1\widetilde{w}_1 + \cdots + b_p\widetilde{w}_p = 0$, forcing all $b_j = 0$ and then all $a_i = 0$.
By 5.39 it now suffices to check that $Tx_k \in \operatorname{span}(x_1, \dots, x_k)$ for each k.
(i) $k = i \leq m$. Applying 5.39 to $T|_U$ gives

$$\begin{aligned} Tu_i = (T|_U)(u_i) &\in \operatorname{span}(u_1, \dots, u_i) \\ &= \operatorname{span}(x_1, \dots, x_i). \end{aligned}$$

(ii) $k = m + j$ with $1 \leq j \leq p$. Applying 5.39 to T/U gives $c_1, \dots, c_j \in \mathbb{F}$ with $(T/U)(\widetilde{w}_j) = c_1\widetilde{w}_1 + \cdots + c_j\widetilde{w}_j$, so

$$\pi(Tw_j) = (T/U)\big(\pi(w_j)\big) = \pi(c_1 w_1 + \cdots + c_j w_j),$$

whence $Tw_j - (c_1 w_1 + \cdots + c_j w_j) \in \operatorname{null} \pi = U$ and therefore

$$\begin{aligned} Tx_{m+j} &\in \operatorname{span}(u_1, \dots, u_m, w_1, \dots, w_j) \\ &= \operatorname{span}(x_1, \dots, x_{m+j}). \end{aligned}$$

Problem 5C.14. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an upper-triangular matrix with respect to some basis of V if and only if the dual operator T' has an upper-triangular matrix with respect to some basis of the dual space V'.

Solution. Each side is equivalent, by 5.44, to the minimal polynomial p of T having the form

$$p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$$

for some $\lambda_1, \dots, \lambda_m \in \mathbb{F}$.
For T itself this is 5.44 applied on V. For T', the space V' is finite-dimensional with $\dim V' = \dim V$ (3.111), so 5.44 applies on V' as well, and the minimal polynomial of T' is again p by Exercise 28 in Section 5B. Hence the two triangularizability statements are equivalent.

5.4 Exercises 5D

Problem 5D.1. Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$.
(a) Prove that if $T^4 = I$, then T is diagonalizable.
(b) Prove that if $T^4 = T$, then T is diagonalizable.
(c) Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^2)$ such that $T^4 = T^2$ and T is not diagonalizable.

Solution. Parts (a) and (b) both follow from this: if $q \in \mathcal{P}(\mathbf{C})$ has no repeated zero and $q(T) = 0$, then T is diagonalizable.
Indeed, the minimal polynomial p of T then divides q by 5.29, say $q = ps$, and 4.13 lets us write

$$p(z) = (z - \alpha_1)^{m_1} \cdots (z - \alpha_k)^{m_k}$$

with $\alpha_1, \ldots, \alpha_k \in \mathbf{C}$ distinct and each $m_j \geq 1$ (with $k = 0$ if $p = 1$). If some $m_j \geq 2$, then $(z - \alpha_j)^2$ divides q, so factoring q into linear factors repeats $z - \alpha_j$, contradicting the uniqueness in 4.13 since the zeros of q are distinct. Hence every $m_j = 1$, and 5.62 makes T diagonalizable.
(a) $T^4 = I$ gives $q(T) = 0$ for $q(z) = z^4 - 1 = (z-1)(z+1)(z-i)(z+i)$, whose zeros $1, -1, i, -i$ are distinct.
(b) $T^4 = T$ gives $q(T) = 0$ for

$$q(z) = z^4 - z = z(z-1)(z-\omega)(z-\omega^2), \qquad \omega = e^{2\pi i/3},$$

whose zeros $0, 1, \omega, \omega^2$ are distinct, since ω is not real, so $\omega^2 = \overline{\omega} \neq \omega$, and neither is 0 or 1.
(c) Take $T(w, z) = (z, 0)$ on $\mathbf{C}^2$. Then $T^2 = 0$, so $T^4 = T^2$, while $T(0,1) = (1,0)$ shows $T \neq 0$; hence the minimal polynomial of T is z^2, which is not a product of distinct linear factors, so T is not diagonalizable by 5.62.

Problem 5D.2. Suppose $T \in \mathcal{L}(V)$ has a diagonal matrix A with respect to some basis of V. Prove that if $\lambda \in \mathbf{F}$, then λ appears on the diagonal of A precisely $\dim E(\lambda, T)$ times.

Solution. Everything follows from the identity $E(\lambda, T) = \operatorname{span}(v_k : k \in S)$, where $v_1, \ldots, v_n$ is the basis giving the diagonal matrix A, the numbers $\lambda_1, \ldots, \lambda_n$ are its diagonal entries, so that $Tv_k = \lambda_k v_k$, and

$$S = \{k \in \{1, \ldots, n\} : \lambda_k = \lambda\}.$$

The inclusion $\supseteq$ holds because $Tv_k = \lambda v_k$ for $k \in S$ and $E(\lambda, T)$ is a subspace (5.52). For $\subseteq$, let $v = a_1 v_1 + \cdots + a_n v_n$ lie in $E(\lambda, T)$; then

$$\begin{aligned} a_1 \lambda_1 v_1 + \cdots + a_n \lambda_n v_n &= Tv = \lambda v \\ &= a_1 \lambda v_1 + \cdots + a_n \lambda v_n, \end{aligned}$$

and uniqueness of the expansion in the basis $v_1, \ldots, v_n$ gives $a_k(\lambda_k - \lambda) = 0$, so $a_k = 0$ whenever $k \notin S$. Being a sublist of a basis, $(v_k)_{k \in S}$ is linearly independent, hence a basis of $E(\lambda, T)$. Therefore $\dim E(\lambda, T)$ is the number of elements of S, that is, the number of times λ appears on the diagonal of A.

Problem 5D.3. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that if the operator T is diagonalizable, then $V = \operatorname{null} T \oplus \operatorname{range} T$.

Solution. Split a basis of eigenvectors according to whether its eigenvalue vanishes: by 5.55(b) there is a basis $v_1, \ldots, v_n$ of V with $Tv_k = \lambda_k v_k$, and with $S = \{k : \lambda_k = 0\}$ and $S' = \{k : \lambda_k \neq 0\}$ we put

$$\begin{aligned} U &= \operatorname{span}(v_k : k \in S), \\ W &= \operatorname{span}(v_k : k \in S'). \end{aligned}$$

Then $V = U \oplus W$: the sum contains every v_k, hence equals V, and if $u + w = 0$ with $u \in U$ and $w \in W$, then expanding both in $v_1, \ldots, v_n$ and using linear independence gives $u = w = 0$, so the sum is direct by 1.45.
Moreover $U = E(0, T) = \operatorname{null} T$, by the identity proved in Exercise 2 of this section taken with $\lambda = 0$, and $W = \operatorname{range} T$ because

$$\begin{aligned} \operatorname{range} T &= \operatorname{span}(Tv_1, \ldots, Tv_n) \\ &= \operatorname{span}(\lambda_1 v_1, \ldots, \lambda_n v_n) = W, \end{aligned}$$

the terms with $k \in S$ being 0 and $\operatorname{span}(\lambda_k v_k) = \operatorname{span}(v_k)$ for $k \in S'$. Hence $V = \operatorname{null} T \oplus \operatorname{range} T$.

Problem 5D.4. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent.
(a) $V = \operatorname{null} T \oplus \operatorname{range} T$.
(b) $V = \operatorname{null} T + \operatorname{range} T$.
(c) $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

Solution. Write $N = \operatorname{null} T$ and $R = \operatorname{range} T$, so that $\dim N + \dim R = \dim V$ by 3.21, and prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).
(a) $\Rightarrow$ (b). The statement $V = N \oplus R$ asserts in particular that $V = N + R$.
(b) $\Rightarrow$ (c). By 2.43,

$$\dim V = \dim(N + R) = \dim N + \dim R - \dim(N \cap R),$$

whose right side is $\dim V - \dim(N \cap R)$ by 3.21. Hence $\dim(N \cap R) = 0$, that is, $N \cap R = \{0\}$.
(c) $\Rightarrow$ (a). The sum $N + R$ is direct by 1.46, and by 2.43 and 3.21,

$$\dim(N \oplus R) = \dim N + \dim R - \dim(N \cap R) = \dim V.$$

A subspace of V with the same dimension as V equals V (2.39), so $V = N \oplus R$.

Problem 5D.5. Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

for every $\lambda \in \mathbf{C}$.

Solution. If T is diagonalizable, a basis $v_1, \dots, v_n$ of eigenvectors of T (5.55(b)) consists of eigenvectors of $T - \lambda I$ as well, since $(T - \lambda I)v_k = (\lambda_k - \lambda)v_k$. So $T - \lambda I$ is diagonalizable by 5.55, and Exercise 3 of this section applied to $T - \lambda I$ gives the displayed decomposition for every λ.
Conversely, assume that decomposition for every $\lambda \in \mathbf{C}$, so 1.46 gives

$$\operatorname{null}(T - \lambda I) \cap \operatorname{range}(T - \lambda I) = \{0\}$$

for every $\lambda \in \mathbf{C}$. Because $\mathbf{F} = \mathbf{C}$, the factorization 4.13 writes the minimal polynomial of T as $p(z) = (z - \lambda_1)^{m_1} \cdots (z - \lambda_k)^{m_k}$ with $\lambda_1, \dots, \lambda_k$ distinct and each $m_j \geq 1$ (with $k = 0$ if $p = 1$). Suppose some $m_j \geq 2$; relabel so $j = 1$ and set

$$r(z) = \frac{p(z)}{(z - \lambda_1)^2}, \qquad q(z) = (z - \lambda_1)r(z),$$

both genuine polynomials since $m_1 \geq 2$. As $q \neq 0$ has degree less than $\deg p$, it is not a polynomial multiple of p, so $q(T) \neq 0$ by 5.29; choose v with $q(T)v \neq 0$ and set $u = (T - \lambda_1 I)r(T)v$. By 5.17,

$$u = q(T)v \neq 0, \qquad (T - \lambda_1 I)u = p(T)v = 0,$$

so u is a nonzero vector of $\operatorname{null}(T - \lambda_1 I) \cap \operatorname{range}(T - \lambda_1 I)$, contradicting the display at $\lambda = \lambda_1$. Hence every $m_j = 1$, and 5.62 makes T diagonalizable.

Problem 5D.6. Suppose $T \in \mathcal{L}(\mathbf{F}^5)$ and $\dim E(8, T) = 4$. Prove that $T - 2I$ or $T - 6I$ is invertible.

Solution. Suppose neither $T - 2I$ nor $T - 6I$ is invertible. Then neither is injective, by 3.65 applied on the finite-dimensional space $\mathbf{F}^5$, so $E(2, T)$ and $E(6, T)$ are nonzero; together with $\dim E(8, T) = 4$ this makes $2, 6, 8$ three distinct eigenvalues of T. Then 5.54 gives

$$6 \leq \dim E(2, T) + \dim E(6, T) + \dim E(8, T) \leq \dim \mathbf{F}^5 = 5,$$

a contradiction. Hence $T - 2I$ or $T - 6I$ is invertible.

Problem 5D.7. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$E(\lambda, T) = E\big(\tfrac{1}{\lambda}, T^{-1}\big)$$

for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

Solution. For $v \in V$ the conditions $Tv = \lambda v$ and $T^{-1}v = \frac{1}{\lambda}v$ are equivalent, which is the assertion, since by 5.52 these conditions describe membership in $E(\lambda, T)$ and in $E\big(\frac{1}{\lambda}, T^{-1}\big)$ respectively.
Applying T^{-1} to $Tv = \lambda v$ gives $v = \lambda T^{-1}v$, hence $T^{-1}v = \frac{1}{\lambda}v$ as $\lambda \neq 0$; applying T to $T^{-1}v = \frac{1}{\lambda}v$ gives $v = \frac{1}{\lambda}Tv$, hence $Tv = \lambda v$.

Problem 5D.8. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct nonzero eigenvalues of T. Prove that

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim \operatorname{range} T.$$

Solution. Each $E(\lambda_k, T)$ lies inside $\operatorname{range} T$, since $Tv = \lambda_k v$ with $\lambda_k \neq 0$ gives $v = T\Big(\frac{1}{\lambda_k}v\Big)$; hence so does their sum $U = E(\lambda_1, T) + \cdots + E(\lambda_m, T)$.
That sum is direct, the λ_k being distinct eigenvalues (5.54), so 3.94 identifies $\dim U$ with the sum of the dimensions and

$$\begin{aligned} \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) &= \dim U \\ &\leq \dim \operatorname{range} T, \end{aligned}$$

the inequality by 2.37, since U is a subspace of the finite-dimensional space $\operatorname{range} T$.

Problem 5D.9. Suppose $R, T \in \mathcal{L}(\mathbf{F}^3)$ each have $2, 6, 7$ as eigenvalues. Prove that there exists an invertible operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $R = S^{-1}TS$.

Solution. Take $S \in \mathcal{L}(\mathbf{F}^3)$ with $Su_k = v_k$ for $k = 1, 2, 3$, where u_1, u_2, u_3 and v_1, v_2, v_3 are bases of $\mathbf{F}^3$ with

$$Ru_k = \lambda_k u_k \qquad \text{and} \qquad Tv_k = \lambda_k v_k$$

for $\lambda_1 = 2$, $\lambda_2 = 6$, $\lambda_3 = 7$.
Such bases exist because R and T each have $3 = \dim \mathbf{F}^3$ distinct eigenvalues, so each is diagonalizable by 5.58; picking an eigenvector for each of $2, 6, 7$ in that order gives the matching labels. The map S exists by 3.4, and its range contains the spanning list v_1, v_2, v_3, so S is surjective and hence invertible by 3.65. Finally,

$$(S^{-1}TS)u_k = S^{-1}Tv_k = \lambda_k S^{-1}v_k = \lambda_k u_k = Ru_k$$

for each k, so $S^{-1}TS$ and R agree on a basis of $\mathbf{F}^3$ and are therefore equal.

Problem 5D.10. Find $R, T \in \mathcal{L}(\mathbf{F}^4)$ such that R and T each have $2, 6, 7$ as eigenvalues, R and T have no other eigenvalues, and there does not exist an invertible operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $R = S^{-1}TS$.

Solution. Define $R, T \in \mathcal{L}(\mathbf{F}^4)$ by

$$R(z_1, z_2, z_3, z_4) = (2z_1,\ 6z_2,\ 7z_3,\ 7z_4),$$

$$T(z_1, z_2, z_3, z_4) = (2z_1,\ 6z_2,\ 7z_3 + z_4,\ 7z_4).$$

With respect to the standard basis e_1, e_2, e_3, e_4 both matrices are upper triangular with diagonal entries $2, 6, 7, 7$, so by 5.41 each of R and T has eigenvalues exactly $2, 6, 7$.
The 7-eigenspaces differ in dimension: $E(7, R) = \operatorname{span}(e_3, e_4)$, whereas $Tz = 7z$ forces $z_1 = z_2 = 0$ (from $2z_1 = 7z_1$ and $6z_2 = 7z_2$) and $z_4 = 0$ (from $7z_3 + z_4 = 7z_3$), so $E(7, T) = \operatorname{span}(e_3)$. Thus $\dim E(7, R) = 2$ and $\dim E(7, T) = 1$.
Suppose S were invertible with $R = S^{-1}TS$, that is, $SR = TS$. For $v \in E(7, R)$,

$$T(Sv) = (TS)v = (SR)v = S(7v) = 7(Sv),$$

so S maps $E(7, R)$ into $E(7, T)$, injectively since S is injective; hence $2 = \dim E(7, R) \leq \dim E(7, T) = 1$ by 3.21, a contradiction. So no such S exists.

Problem 5D.11. Find $T \in \mathcal{L}(\mathbf{C}^3)$ such that 6 and 7 are eigenvalues of T and such that T does not have a diagonal matrix with respect to any basis of $\mathbf{C}^3$.

Solution. Define $T \in \mathcal{L}(\mathbf{C}^3)$ by

$$T(z_1, z_2, z_3) = (6z_1 + z_2,\ 6z_2,\ 7z_3).$$

Its matrix with respect to the standard basis e_1, e_2, e_3 is upper triangular with diagonal entries $6, 6, 7$, so by 5.41 the eigenvalues of T are exactly 6 and 7.
Both eigenspaces are one-dimensional: $Tz = 6z$ reads $6z_1 + z_2 = 6z_1$, $7z_3 = 6z_3$, forcing $z_2 = z_3 = 0$, so $E(6, T) = \operatorname{span}(e_1)$; and $Tz = 7z$ reads $6z_2 = 7z_2$, $6z_1 + z_2 = 7z_1$, forcing $z_2 = 0$ and then $z_1 = 0$, so $E(7, T) = \operatorname{span}(e_3)$. Hence

$$\dim E(6, T) + \dim E(7, T) = 1 + 1 = 2 < 3 = \dim \mathbf{C}^3,$$

so condition (d) of 5.55 fails and T is diagonalizable with respect to no basis of $\mathbf{C}^3$.
Method (2): were T diagonalizable, 5.62 together with 5.27 (the zeros of the minimal polynomial are the eigenvalues $6, 7$) would make the minimal polynomial $(z - 6)(z - 7)$. But $Te_2 = e_1 + 6e_2$ gives

$$(T - 6I)(T - 7I)e_2 = (T - 6I)(e_1 - e_2) = 0 - e_1 = -e_1 \neq 0.$$

Problem 5D.12. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is such that 6 and 7 are eigenvalues of T. Furthermore, suppose T does not have a diagonal matrix with respect to any basis of $\mathbf{C}^3$. Prove that there exists $(z_1, z_2, z_3) \in \mathbf{C}^3$ such that

$$T(z_1, z_2, z_3) = (6 + 8z_1,\ 7 + 8z_2,\ 13 + 8z_3).$$

Solution. The desired equation says exactly that $(T - 8I)z = (6, 7, 13)$, so it suffices to show that $T - 8I$ is surjective.
Were 8 an eigenvalue of T, then $6, 7, 8$ would be $3 = \dim \mathbf{C}^3$ distinct eigenvalues of T, making T diagonalizable by 5.58 and contradicting the hypothesis. Hence $\operatorname{null}(T - 8I) = \{0\}$, so $T - 8I$ is injective (3.15) and therefore surjective, injectivity and surjectivity being equivalent for an operator on the finite-dimensional space $\mathbf{C}^3$ (3.65). Thus some $z = (z_1, z_2, z_3)$ satisfies $(T - 8I)z = (6, 7, 13)$, which is to say

$$T(z_1, z_2, z_3) = (6 + 8z_1,\ 7 + 8z_2,\ 13 + 8z_3).$$

Problem 5D.13. Suppose A is a diagonal matrix with distinct entries on the diagonal and B is a matrix of the same size as A. Show that $AB = BA$ if and only if B is a diagonal matrix.

Solution. Say A, B are n-by-n with $A_{j,j} = \lambda_j$ and $A_{j,k} = 0$ for $j \neq k$, the λ_j distinct. Because all but one term of each sum vanishes, the entry formula 3.47 gives

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r}B_{r,k} = \lambda_j B_{j,k}, \qquad (BA)_{j,k} = \sum_{r=1}^{n} B_{j,r}A_{r,k} = \lambda_k B_{j,k},$$

so, comparing entries,

$$AB = BA \iff (\lambda_j - \lambda_k)B_{j,k} = 0 \text{ for all } j, k.$$

(i) If $AB = BA$ and $j \neq k$, then $\lambda_j - \lambda_k \neq 0$ by distinctness, so $B_{j,k} = 0$; hence B is diagonal.
(ii) If B is diagonal, then $B_{j,k} = 0$ when $j \neq k$ and $\lambda_j - \lambda_k = 0$ when $j = k$, so the displayed condition holds and $AB = BA$.

Problem 5D.14. (a) Give an example of a finite-dimensional complex vector space and an operator T on that vector space such that T^2 is diagonalizable but T is not diagonalizable.
(b) Suppose $\mathbf{F} = \mathbf{C}$, k is a positive integer, and $T \in \mathcal{L}(V)$ is invertible. Prove that T is diagonalizable if and only if T^k is diagonalizable.

Solution. (a) Take $V = \mathbf{C}^2$ and define $T \in \mathcal{L}(\mathbf{C}^2)$ by

$$T(z_1, z_2) = (z_2, 0).$$

Then $T^2 = 0$, whose matrix is diagonal with respect to every basis. But the matrix of T with respect to the standard basis is upper triangular with both diagonal entries 0, so 0 is the only eigenvalue of T by 5.41, and

$$\begin{aligned} E(0, T) = \operatorname{null} T &= \{(z_1, z_2) \in \mathbf{C}^2 : z_2 = 0\} \\ &= \operatorname{span}\big((1, 0)\big) \end{aligned}$$

has dimension $1 < 2 = \dim \mathbf{C}^2$; so condition (d) of 5.55 fails and T is not diagonalizable.
(b) If T is diagonalizable, take a basis $v_1, \ldots, v_n$ of eigenvectors, $Tv_j = \lambda_j v_j$; then $T^k v_j = \lambda_j^k v_j$, so the same basis diagonalizes T^k (5.55(b)).
Conversely, suppose T^k is diagonalizable. By 5.62 the minimal polynomial of T^k is $(z - \mu_1) \cdots (z - \mu_m)$ with $\mu_1, \ldots, \mu_m \in \mathbf{C}$ distinct, and each $\mu_j \neq 0$, since T^k is invertible and the zeros of the minimal polynomial are the eigenvalues (5.27). Set

$$p(z) = (z^k - \mu_1)(z^k - \mu_2) \cdots (z^k - \mu_m).$$

Multiplicativity of $q \mapsto q(T)$ (5.17) gives

$$p(T) = (T^k - \mu_1 I) \cdots (T^k - \mu_m I) = 0,$$

this being the minimal polynomial of T^k evaluated at T^k. Each nonzero μ_j has exactly k distinct k-th roots in $\mathbf{C}$, and a zero shared by $z^k - \mu_j$ and $z^k - \mu_{j'}$ would force $\mu_j = \mu_{j'}$; so p, monic of degree km, has km distinct zeros and its factorization 4.13 is into km distinct linear factors. Hence no $(z - \alpha)^2$ divides p.
Let q be the minimal polynomial of T, so $p = qs$ for some $s \in \mathcal{P}(\mathbf{C})$ by 5.29. Since $\mathbf{F} = \mathbf{C}$, 4.13 gives $q = (z - \lambda_1)^{d_1} \cdots (z - \lambda_r)^{d_r}$ with $\lambda_1, \ldots, \lambda_r$ distinct. Any $d_i \geq 2$ would make $(z - \lambda_i)^2$ a divisor of q and hence of p, which is impossible; so every $d_i = 1$ and T is diagonalizable by 5.62.

Problem 5D.15. Suppose V is a finite-dimensional complex vector space, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Prove that the following are equivalent.
(a) T is diagonalizable.
(b) There does not exist $\lambda \in \mathbf{C}$ such that p is a polynomial multiple of $(z-\lambda)^2$.
(c) p and its derivative p' have no zeros in common.
(d) The greatest common divisor of p and p' is the constant polynomial 1.
The greatest common divisor of p and p' is the monic polynomial q of largest degree such that p and p' are both polynomial multiples of q. The Euclidean algorithm for polynomials (look it up) can quickly determine the greatest common divisor of two polynomials, without requiring any information about the zeros of the polynomials. Thus the equivalence of (a) and (d) above shows that we can determine whether T is diagonalizable without knowing anything about the zeros of p.

Solution. We prove (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) $\Leftrightarrow$ (d), each by contraposition. (If $V = \{0\}$ then $p = 1$ and all four statements hold trivially, so assume $V \neq \{0\}$, whence $\deg p \geq 1$.)
(a) $\Leftrightarrow$ (b). If (b) fails, say $p = (z-\lambda)^2 r$, and T were diagonalizable, then 5.62 would give $p = (z-\mu_1)\cdots(z-\mu_m)$ with $\mu_1, \ldots, \mu_m$ distinct; from $p(\lambda) = 0$ we get $\lambda = \mu_j$ for exactly one j, so $p = (z-\lambda)s$ with $s = \prod_{k\neq j}(z-\mu_k)$ satisfying $s(\lambda) \neq 0$. Comparing with $p = (z-\lambda)\big((z-\lambda)r\big)$ and cancelling $z-\lambda$ (legitimate since $\mathcal{P}(\mathbf{C})$ has no zero divisors) gives $s = (z-\lambda)r$ and hence $s(\lambda) = 0$, a contradiction; so (a) fails. Conversely, if (b) holds then 5.27(b) writes $p = (z-\lambda_1)\cdots(z-\lambda_n)$ over $\mathbf{C}$, and a repetition $\lambda_j = \lambda_k$ with $j \neq k$ would exhibit p as a multiple of $(z-\lambda_j)^2$; so the λ_i are distinct and 5.62 makes T diagonalizable.
(b) $\Leftrightarrow$ (c). If $p = (z-\lambda)^2 r$, then

$$p' = 2(z-\lambda)r + (z-\lambda)^2 r',$$

so $p(\lambda) = p'(\lambda) = 0$ and (c) fails. Conversely, if $p(\lambda) = p'(\lambda) = 0$, the factor theorem 4.6 gives $p = (z-\lambda)q$, whence $p' = q + (z-\lambda)q'$ and $q(\lambda) = p'(\lambda) = 0$; as $q \neq 0$ (because $p \neq 0$), 4.6 applied to q gives $q = (z-\lambda)r$ and so $p = (z-\lambda)^2 r$, so (b) fails.
(c) $\Leftrightarrow$ (d). Here $p' \neq 0$, its leading coefficient being $\deg p \neq 0$ in $\mathbf{C}$, so the greatest common divisor exists. If $p(\lambda) = p'(\lambda) = 0$, then $z-\lambda$ divides both p and p' by 4.6, so their greatest common divisor has degree at least 1 and (d) fails. Conversely, if that greatest common divisor q has $\deg q \geq 1$, then 4.12 gives it a zero λ, and p, p' being multiples of q yields $p(\lambda) = p'(\lambda) = 0$, so (c) fails.

Problem 5D.16. Suppose that $T \in \mathcal{L}(V)$ is diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Prove that a subspace U of V is invariant under T if and only if there exist subspaces $U_1, \ldots, U_m$ of V such that $U_k \subseteq E(\lambda_k, T)$ for each k and $U = U_1 \oplus \cdots \oplus U_m$.

Solution. In the forward direction the subspaces to take are $U_k = U \cap E(\lambda_k, T)$. Note V is finite-dimensional and, by 5.55(c),

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T).$$

(i) Suppose $U_k \subseteq E(\lambda_k, T)$ for each k and $U = U_1 \oplus \cdots \oplus U_m$. Writing $u \in U$ as $u = u_1 + \cdots + u_m$ with $u_k \in U_k$, each $Tu_k = \lambda_k u_k$ lies in the subspace U_k, so

$$Tu = \lambda_1 u_1 + \cdots + \lambda_m u_m \in U_1 + \cdots + U_m = U,$$

and U is invariant under T.
(ii) Suppose U is invariant under T and put $U_k = U \cap E(\lambda_k, T)$, a subspace of $E(\lambda_k, T)$ contained in U. The sum $U_1 + \cdots + U_m$ is direct: if $u_1 + \cdots + u_m = 0$ with $u_k \in U_k \subseteq E(\lambda_k, T)$, then each $u_k = 0$ because the sum of the eigenspaces is direct (5.54), so 1.45 applies. Since U is invariant under the diagonalizable T, the restriction $T|_U$ is diagonalizable (5.65) on the finite-dimensional space U (2.25), so by 5.55(b) U has a basis of eigenvectors of $T|_U$; each such vector is an eigenvector of T, hence lies in $E(\lambda_k, T)$ for some k and so in U_k. Therefore

$$U \subseteq U_1 + \cdots + U_m \subseteq U,$$

giving $U = U_1 \oplus \cdots \oplus U_m$.

Problem 5D.17. Suppose V is finite-dimensional. Prove that $\mathcal{L}(V)$ has a basis consisting of diagonalizable operators.

Solution. Fix a basis $v_1, \ldots, v_n$ of V and let $E_{j,k} \in \mathcal{L}(V)$ be the operator with $E_{j,k}v_k = v_j$ and $E_{j,k}v_l = 0$ for $l \neq k$ (3.4); then the list of n^2 operators

$$E_{j,j} \quad (j = 1, \ldots, n) \qquad \text{and} \qquad E_{j,j} + E_{j,k} \quad (j \neq k)$$

is a basis of $\mathcal{L}(V)$ consisting of diagonalizable operators. (For $n = 0$ the empty list serves.)
It is a basis: the $E_{j,k}$ have as matrices the n^2 standard matrix units, which form a basis of $\mathbf{F}^{n,n}$, and $\mathcal{M}$ is an isomorphism of $\mathcal{L}(V)$ onto $\mathbf{F}^{n,n}$ (3.71), so the $E_{j,k}$ form a basis of $\mathcal{L}(V)$ and $\dim \mathcal{L}(V) = n^2$; our list spans them all, since $E_{j,k} = (E_{j,j} + E_{j,k}) - E_{j,j}$ for $j \neq k$, and a spanning list of length n^2 is a basis (2.42).
Each $E_{j,j}$ is diagonalizable, since $v_1, \ldots, v_n$ is a basis of eigenvectors of it (eigenvalue 1 for v_j, eigenvalue 0 for the rest), so 5.55(b) applies. For $j \neq k$ put $S = E_{j,j} + E_{j,k}$, so that $Sv_j = v_j$, $Sv_k = v_j$, and $Sv_l = 0$ for $l \notin \{j, k\}$. Replacing v_k by $v_k - v_j$ gives a list of length n still spanning V, hence a basis (2.42), and it consists of eigenvectors of S:

$$Sv_j = 1 \cdot v_j, \qquad S(v_k - v_j) = 0 \cdot (v_k - v_j), \qquad Sv_l = 0 \cdot v_l.$$

So S too is diagonalizable by 5.55(b).

Problem 5D.18. Suppose that $T \in \mathcal{L}(V)$ is diagonalizable and U is a subspace of V that is invariant under T. Prove that the quotient operator T/U is a diagonalizable operator on V/U.
The quotient operator T/U was defined in Exercise 38 in Section 5A.

Solution. Push a basis of eigenvectors of T down to V/U. Since T is diagonalizable, V is finite-dimensional and 5.55(b) gives a basis $v_1, \ldots, v_n$ of V with $Tv_i = \lambda_i v_i$. Let $\pi\colon V \to V/U$ be the quotient map $\pi v = v + U$, which is linear and surjective and satisfies $(T/U)(\pi v) = Tv + U = \pi(Tv)$ by the definition of the quotient operator (Exercise 38 in Section 5A). Hence

$$(T/U)(\pi v_i) = \pi(Tv_i) = \lambda_i \, \pi v_i,$$

so every nonzero πv_i is an eigenvector of T/U; and $\pi v_1, \ldots, \pi v_n$ spans V/U because π is linear and surjective. Discarding the zero entries leaves a spanning list of eigenvectors of T/U, which contains a basis of V/U (2.30). By 5.55(b), T/U is diagonalizable.

Problem 5D.19. Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there exists a subspace U of V that is invariant under T such that $T|_U$ and T/U are both diagonalizable, then T is diagonalizable.
See Exercise 13 in Section 5C for an analogous statement about upper-triangular matrices.

Solution. False: take $V = \mathbf{F}^2$ with $T(x, y) = (y, 0)$ and $U = \operatorname{span}\bigl((1, 0)\bigr)$.
Then U is invariant under T, since $T(x, 0) = (0, 0)$, and $T|_U = 0$ is diagonalizable on the one-dimensional space U. Also $T/U = 0$ on V/U, since for every $(x, y) \in \mathbf{F}^2$,

$$(T/U)\bigl((x, y) + U\bigr) = (y, 0) + U = 0 + U,$$

so T/U is diagonalizable as well.
But $T \neq 0$ and $T^2 = 0$, so the minimal polynomial of T is z^2, which is not a product of distinct linear factors; hence T is not diagonalizable by 5.62.

Problem 5D.20. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if the dual operator T' is diagonalizable.

Solution. The operators T and T' have the same minimal polynomial, after which 5.62 settles both directions at once.
First, $\big(q(T)\big)' = q(T')$ for every $q \in \mathcal{P}(\mathbf{F})$. Indeed $I' = I$, since $I'\varphi = \varphi \circ I = \varphi$, and $(T^k)' = (T')^k$ by induction from 3.120(c):

$$\begin{aligned}(T^k)' &= (T\,T^{k-1})' = (T^{k-1})'\,T' \\ &= (T')^{k-1}T' = (T')^k.\end{aligned}$$

So for $q(z) = a_0 + a_1 z + \cdots + a_d z^d$, linearity of $S \mapsto S'$ (3.120(a) and (b)) gives

$$\big(q(T)\big)' = a_0 I + a_1 T' + \cdots + a_d (T')^d = q(T').$$

Second, $S' = 0$ forces $S = 0$: if $Sv \neq 0$, extend the linearly independent list Sv to a basis of V (2.32) and use 3.4 to get $\varphi \in V'$ with $\varphi(Sv) = 1$ and φ zero on the other basis vectors; then $(S'\varphi)(v) = \varphi(Sv) = 1 \neq 0$.
Combining, $q(T) = 0 \iff q(T') = 0$ for every q, so the monic polynomial of smallest degree annihilating T is the one annihilating T' (5.24); call it p. Since $\dim V' = \dim V < \infty$ (3.111), 5.62 applies to each of T and T': each is diagonalizable exactly when p is a product of distinct linear factors. Hence T is diagonalizable if and only if T' is.

Problem 5D.21. The Fibonacci sequence $F_0, F_1, F_2, \ldots$ is defined by

$$F_0 = 0, \quad F_1 = 1, \quad \text{and} \quad F_n = F_{n-2} + F_{n-1} \text{ for } n \geq 2.$$

Define $T \in \mathcal{L}(\mathbf{R}^2)$ by $T(x, y) = (y, x + y)$.
(a) Show that $T^n(0, 1) = (F_n, F_{n+1})$ for each nonnegative integer n.
(b) Find the eigenvalues of T.
(c) Find a basis of $\mathbf{R}^2$ consisting of eigenvectors of T.
(d) Use the solution to (c) to compute $T^n(0, 1)$. Conclude that

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

for each nonnegative integer n.
(e) Use (d) to conclude that if n is a nonnegative integer, then the Fibonacci number F_n is the integer that is closest to

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

Each F_n is a nonnegative integer, even though the right side of the formula in (d) does not look like an integer. The number

$$\frac{1+\sqrt{5}}{2}$$

is called the golden ratio.

Solution. Throughout, write

$$\varphi = \frac{1+\sqrt{5}}{2}, \qquad \psi = \frac{1-\sqrt{5}}{2}.$$

Note $\varphi - \psi = \sqrt{5}$ and $\lambda^2 = \lambda + 1$ for $\lambda \in \{\varphi, \psi\}$.
(a) Induct on n. For $n = 0$, $T^0(0, 1) = (0, 1) = (F_0, F_1)$; and if $T^n(0, 1) = (F_n, F_{n+1})$, then the

recurrence gives

$$T^{n+1}(0,1) = T(F_n, F_{n+1}) = (F_{n+1},\, F_n + F_{n+1}) = (F_{n+1}, F_{n+2}).$$

(b) The eigenvalues are φ and ψ. Indeed $T(x,y) = \lambda(x,y)$ says $y = \lambda x$ and $x+y = \lambda y$, so $(\lambda^2-\lambda-1)x = 0$; here $x = 0$ would force $y = \lambda x = 0$, so $\lambda^2 - \lambda - 1 = 0$ and $\lambda = \frac{1\pm\sqrt{5}}{2}$. Conversely each such λ is an eigenvalue, since

$$T(1,\lambda) = (\lambda, 1+\lambda) = (\lambda, \lambda^2) = \lambda(1,\lambda).$$

(c) The basis is $(1,\varphi), (1,\psi)$: by (b) these are eigenvectors for the distinct eigenvalues $\varphi \neq \psi$, hence linearly independent (5.11), and a linearly independent list of length $2 = \dim \mathbf{R}^2$ is a basis (2.38).
(d) Writing $(0,1) = a(1,\varphi)+b(1,\psi)$ gives $a+b = 0$ and $a\varphi+b\psi = 1$, so $a(\varphi-\psi) = 1$ and $a = 1/\sqrt{5} = -b$. Applying T^n,

$$T^n(0,1) = \frac{\varphi^n}{\sqrt{5}}(1,\varphi) - \frac{\psi^n}{\sqrt{5}}(1,\psi),$$

whose first coordinate is $(\varphi^n - \psi^n)/\sqrt{5}$ and equals F_n by (a). That is the asserted formula.
(e) Since $2 < \sqrt{5} < 3$ we have $-1 < \psi < 0$, so $|\psi|^n \leq 1$ and (d) gives

$$\left| F_n - \frac{1}{\sqrt{5}}\varphi^n \right| = \frac{|\psi|^n}{\sqrt{5}} \leq \frac{1}{\sqrt{5}} < \frac{1}{2}.$$

An integer m with $|t - m| < \frac{1}{2}$ is the unique closest integer to t, since any other $m' \in \mathbf{Z}$ has $|t - m'| \geq |m - m'| - |t - m| > \frac{1}{2} > |t - m|$. Hence F_n is the integer closest to $\varphi^n/\sqrt{5}$.

Problem 5D.22. Suppose $T \in \mathcal{L}(V)$ and A is an n-by-n matrix that is the matrix of T with respect to some basis of V. Prove that if

$$|A_{j,j}| > \sum_{\substack{k=1 \\ k\neq j}}^{n} |A_{j,k}|$$

for each $j \in \{1,\ldots,n\}$, then T is invertible.
This exercise states that if the diagonal entries of the matrix of T are large compared to the nondiagonal entries, then T is invertible.

Solution. Apply the Gershgorin disk theorem 5.67 to the basis $v_1,\ldots,v_n$ of V with respect to which $A = \mathcal{M}(T)$; note $\dim V = n < \infty$, so T is invertible if and only if it is injective (3.63 with 3.65), that is, exactly when 0 is not an eigenvalue of T.
If T were not invertible, then 0 would be an eigenvalue of T and hence, by 5.67, lie in some Gershgorin disk: for some j,

$$|A_{j,j}| = |0 - A_{j,j}| \leq \sum_{\substack{k=1 \\ k\neq j}}^{n} |A_{j,k}|,$$

contradicting the strict inequality assumed for every j. Hence T is invertible.

Problem 5D.23. Suppose the definition of the Gershgorin disks is changed so that the radius of the k^{th} disk is the sum of the absolute values of the entries in column (instead of row) k of A, excluding the diagonal entry. Show that the Gershgorin disk theorem (5.67) still holds with this changed definition.

Solution. Write A for the matrix of T with respect to a basis $v_1,\ldots,v_n$ of V; the changed k^{th} disk is

$$D_k = \{z \in \mathbf{F} : |z - A_{k,k}| \leq r_k\}, \qquad r_k = \sum_{\substack{j=1 \\ j\neq k}}^{n} |A_{j,k}|,$$

and the claim is that every eigenvalue of T lies in some D_k.
Let $Tw = \lambda w$ with $w = c_1 v_1 + \cdots + c_n v_n \neq 0$. Exactly as in the proof of 5.67, expanding $Tv_k = \sum_j A_{j,k} v_j$ and using uniqueness of the basis representation of λw gives $\lambda c_j = \sum_k A_{j,k} c_k$ for each j, hence by the triangle inequality

$$|\lambda - A_{j,j}|\,|c_j| \leq \sum_{\substack{k=1\\k\neq j}}^{n} |A_{j,k}|\,|c_k|.$$

Rather than single out a largest $|c_j|$, sum over j and interchange the order of summation in the resulting finite double sum over $\{(j,k) : j \neq k\}$:

$$\begin{aligned}\sum_{j=1}^{n} |\lambda - A_{j,j}|\,|c_j| &\leq \sum_{k=1}^{n} |c_k| \sum_{\substack{j=1\\j\neq k}}^{n} |A_{j,k}|\\ &= \sum_{k=1}^{n} |c_k|\, r_k.\end{aligned}$$

If λ lay in no D_k, then $|\lambda - A_{k,k}| > r_k$ for every k; multiplying by $|c_k| \geq 0$ and summing gives a strict inequality, since $c_k \neq 0$ for at least one k, and hence

$$\begin{aligned}\sum_{k=1}^{n} |c_k|\,|\lambda - A_{k,k}| &\leq \sum_{k=1}^{n} |c_k|\, r_k\\ &< \sum_{k=1}^{n} |c_k|\,|\lambda - A_{k,k}|,\end{aligned}$$

a finite number strictly less than itself. Hence $\lambda \in D_k$ for some k.

5.5 Exercises 5E

Problem 5E.1. Give an example of two commuting operators S, T on $\mathbf{F}^4$ such that there is a subspace of $\mathbf{F}^4$ that is invariant under S but not under T and there is a subspace of $\mathbf{F}^4$ that is invariant under T but not under S.

Solution. With e_1, e_2, e_3, e_4 the standard basis of $\mathbf{F}^4$, take

$$S(x_1, x_2, x_3, x_4) = (x_2, 0, 0, 0), \qquad T(x_1, x_2, x_3, x_4) = (0, 0, x_4, 0),$$

so $Se_2 = e_1$ and $Te_4 = e_3$, while each kills the remaining standard basis vectors.
These commute, since $ST = 0 = TS$: the second coordinate of $T(x_1, x_2, x_3, x_4) = (0, 0, x_4, 0)$ is 0, and the fourth coordinate of $S(x_1, x_2, x_3, x_4) = (x_2, 0, 0, 0)$ is 0.
(i) $U = \operatorname{span}(e_1, e_4)$ is invariant under S, as $Se_1 = Se_4 = 0$, but not under T, as $Te_4 = e_3 \notin U$.
(ii) $W = \operatorname{span}(e_2, e_3)$ is invariant under T, as $Te_2 = Te_3 = 0$, but not under S, as $Se_2 = e_1 \notin W$.

Problem 5E.2. Suppose $\mathcal{E}$ is a subset of $\mathcal{L}(V)$ and every element of $\mathcal{E}$ is diagonalizable. Prove that there exists a basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if every pair of elements of $\mathcal{E}$ commutes.
[This exercise extends 5.76, which considers the case in which $\mathcal{E}$ contains only two elements. For this exercise, $\mathcal{E}$ may contain any number of elements, and $\mathcal{E}$ may even be an infinite set.]

Solution. Here V is finite-dimensional, since a diagonalizable operator has a matrix with respect to a (finite) basis.
($\Rightarrow$) If every element of $\mathcal{E}$ has a diagonal matrix with respect to one basis, then for $S, T \in \mathcal{E}$ the matrices $\mathcal{M}(S)$ and $\mathcal{M}(T)$ are diagonal of the same size and so commute, whence $ST = TS$ by 5.74.

($\Leftarrow$) Suppose every pair in $\mathcal{E}$ commutes. We show by strong induction on $\dim V$ that V has a basis of vectors that are eigenvectors of every element of $\mathcal{E}$; with respect to such a basis each $T \in \mathcal{E}$ has a diagonal matrix, since $Tv_k = \mu_k v_k$ makes the k^{th} column μ_k times the k^{th} standard column. If $\dim V \leq 1$, every operator on V is a scalar multiple of I and any basis works. So let $\dim V = n \geq 2$.
(i) If every element of $\mathcal{E}$ is a scalar multiple of I, any basis of V consists of common eigenvectors.
(ii) Otherwise choose $S \in \mathcal{E}$ that is not a scalar multiple of I, with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$. Since S is diagonalizable, 5.55(c) gives

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S),$$

where $m \geq 2$ (else $V = E(\lambda_1, S)$, i.e. $S = \lambda_1 I$); the summands being nonzero, each $\dim E(\lambda_k, S) < n$. Fix k and set $U = E(\lambda_k, S)$. Every $T \in \mathcal{E}$ commutes with S, so U is invariant under T (5.75) and $T|_U$ is diagonalizable (5.65); any two such restrictions commute, since

$$(T|_U)(R|_U)u = (TR)u = (RT)u = (R|_U)(T|_U)u \qquad (u \in U).$$

By the induction hypothesis applied to U and $\{T|_U : T \in \mathcal{E}\}$, the eigenspace $E(\lambda_k, S)$ has a basis of common eigenvectors of the restrictions, and each such vector u is an eigenvector of every $T \in \mathcal{E}$ itself, as $u \neq 0$ and $Tu = (T|_U)u$ is a scalar multiple of u.
Concatenating these bases over $k = 1, \ldots, m$ gives a list spanning V of length $\sum_k \dim E(\lambda_k, S) = \dim V$, hence a basis of V, consisting of common eigenvectors of every element of $\mathcal{E}$. The induction is on $\dim V$, so $\mathcal{E}$ may be infinite.

Problem 5E.3. Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Suppose $p \in \mathcal{P}(\mathbf{F})$.
(a) Prove that $\operatorname{null} p(S)$ is invariant under T.
(b) Prove that $\operatorname{range} p(S)$ is invariant under T.
[See 5.18 for the special case $S = T$.]

Solution. Both parts follow from $p(S)T = Tp(S)$. Indeed $S^k T = TS^k$ for every $k \geq 0$, by induction from $ST = TS$:

$$S^{k+1}T = S(S^k T) = S(TS^k) = (ST)S^k = (TS)S^k = TS^{k+1},$$

so writing $p(S) = a_0 I + a_1 S + \cdots + a_N S^N$ gives

$$p(S)T = \sum_{k=0}^{N} a_k S^k T = \sum_{k=0}^{N} a_k T S^k = Tp(S).$$

(a) If $p(S)v = 0$, then $p(S)(Tv) = \big(Tp(S)\big)v = T0 = 0$, so $Tv \in \operatorname{null} p(S)$.
(b) If $v = p(S)u$, then $Tv = \big(Tp(S)\big)u = p(S)(Tu) \in \operatorname{range} p(S)$.

Problem 5E.4. Prove or give a counterexample: If A is a diagonal matrix and B is an upper-triangular matrix of the same size as A, then A and B commute.

Solution. The statement is *false*. Here is a counterexample of size 2-by-2:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Here A is diagonal and B is upper triangular of the same size, yet

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

so $AB \neq BA$.

Problem 5E.5. Prove that a pair of operators on a finite-dimensional vector space commute if and only if their dual operators commute.
[See 3.118 for the definition of the dual of an operator.]

Solution. Both directions rest on the order reversal $(RT)' = T'R'$ of 3.120(c). Let $S, T \in \mathcal{L}(V)$ with V finite-dimensional.
If $ST = TS$, then

$$T'S' = (ST)' = (TS)' = S'T'.$$

Conversely, if $S'T' = T'S'$, then $(TS)' = (ST)'$, so $R = ST - TS$ satisfies $R' = 0$ by linearity of $R \mapsto R'$ (3.120(a) and (b)). Hence $(\operatorname{range} R)^0 = \operatorname{null} R' = V'$ by 3.128(a), and 3.127(b), applicable since V is finite-dimensional, gives $\operatorname{range} R = \{0\}$. Thus $R = 0$, that is, $ST = TS$.
Method (2): with $A = \mathcal{M}(S)$, $B = \mathcal{M}(T)$ with respect to a basis of V and the dual basis, 3.132 makes the matrices of S', T' equal to $A^{\mathrm{t}}, B^{\mathrm{t}}$, so by 5.74 twice, $(XY)^{\mathrm{t}} = Y^{\mathrm{t}}X^{\mathrm{t}}$, and injectivity of transposition,

$$S'T' = T'S' \iff (BA)^{\mathrm{t}} = (AB)^{\mathrm{t}} \iff BA = AB \iff ST = TS.$$

Problem 5E.6. Suppose that V is a nonzero finite-dimensional complex vector space and $S, T \in \mathcal{L}(V)$ commute. Prove that there exist $\alpha, \lambda \in \mathbf{C}$ such that

$$\operatorname{range}(S - \alpha I) + \operatorname{range}(T - \lambda I) \neq V.$$

Solution. Take $\alpha = A_{n,n}$ and $\lambda = B_{n,n}$, where (using that V is a nonzero finite-dimensional complex vector space and S, T commute) 5.80 supplies a basis $v_1, \ldots, v_n$ of V making both $A = \mathcal{M}(S)$ and $B = \mathcal{M}(T)$ upper triangular. Both ranges then lie in $U = \operatorname{span}(v_1, \ldots, v_{n-1})$, a subspace with $\dim U = n - 1 < n$.
Indeed, for $k \leq n-1$, upper-triangularity of A gives $Sv_k \in \operatorname{span}(v_1, \ldots, v_k) \subseteq U$ by 5.39, and $\alpha v_k \in U$, so $(S - \alpha I)v_k \in U$; while for $k = n$,

$$(S - \alpha I)v_n = \sum_{j=1}^{n} A_{j,n} v_j - A_{n,n} v_n = \sum_{j=1}^{n-1} A_{j,n} v_j \in U.$$

Since $v_1, \ldots, v_n$ spans V, the vectors $(S - \alpha I)v_k$ span $\operatorname{range}(S - \alpha I)$, so that range lies in U; the identical argument with B, λ puts $\operatorname{range}(T - \lambda I)$ in U. Hence

$$\operatorname{range}(S - \alpha I) + \operatorname{range}(T - \lambda I) \subseteq U \neq V.$$

Problem 5E.7. Suppose V is a complex vector space, $S \in \mathcal{L}(V)$ is diagonalizable, and $T \in \mathcal{L}(V)$ commutes with S. Prove that there is a basis of V such that S has a diagonal matrix with respect to this basis and T has an upper-triangular matrix with respect to this basis.

Solution. Concatenate, over the distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ of S, bases of the eigenspaces $E(\lambda_k, S)$ with respect to which T is upper triangular there.
Since S is diagonalizable, V is finite-dimensional and 5.55(c) gives

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S).$$

As S and T commute, each $E(\lambda_k, S)$ is invariant under T by 5.75, so $T|_{E(\lambda_k,S)}$ is an operator on a finite-dimensional complex vector space and 5.47 gives a basis $u_1^{(k)}, \ldots, u_{d_k}^{(k)}$ of $E(\lambda_k, S)$ making it upper triangular; equivalently (5.39),

$$Tu_j^{(k)} \in \operatorname{span}\big(u_1^{(k)}, \ldots, u_j^{(k)}\big) \qquad (1 \leq j \leq d_k).$$

Concatenating these lists in the order $k = 1, \ldots, m$ gives a list $v_1, \ldots, v_n$ spanning V of length $\sum_k \dim E(\lambda_k, S) = \dim V$, hence a basis of V.

Each v_i lies in some $E(\lambda_k, S)$, so $Sv_i = \lambda_k v_i$ and $\mathcal{M}(S,(v_1,\ldots,v_n))$ is diagonal. And if $v_i = u_j^{(k)}$, then $u_1^{(k)},\ldots,u_j^{(k)}$ occupy positions at most i in the concatenation, so the display above gives $Tv_i \in \operatorname{span}(v_1,\ldots,v_i)$; by 5.39 the matrix of T is upper triangular.

Problem 5E.8. Suppose $m = 3$ in Example 5.72 and D_w, D_z are the commuting partial differentiation operators on $\mathcal{P}_3(\mathbf{C}^2,\mathbf{C})$ from that example. Find a basis of $\mathcal{P}_3(\mathbf{C}^2,\mathbf{C})$ with respect to which D_w and D_z each have an upper-triangular matrix.

Solution. Take the monomials $w^j z^k$ with $j+k \le 3$, listed in order of increasing total degree:

$$v_1 = 1;\quad v_2 = w,\ v_3 = z;\quad v_4 = w^2,\ v_5 = wz,\ v_6 = z^2;$$
$$v_7 = w^3,\ v_8 = w^2 z,\ v_9 = wz^2,\ v_{10} = z^3.$$

These span $\mathcal{P}_3(\mathbf{C}^2,\mathbf{C})$ by 5.73 and are linearly independent, since a polynomial in two variables vanishing on all of $\mathbf{C}^2$ has every coefficient 0 (Check!); so they form a basis.
Both operators strictly lower total degree: for $j+k \le 3$,

$$D_w(w^j z^k) = j\, w^{j-1} z^k, \qquad D_z(w^j z^k) = k\, w^j z^{k-1},$$

each a scalar multiple of a monomial of total degree $j+k-1$, or 0 (when $j=0$, respectively $k=0$). Since the list places every monomial of total degree $d-1$ before every monomial of total degree d,

$$D_w v_i,\ D_z v_i \in \operatorname{span}(v_1,\ldots,v_{i-1})$$

for each i, which lies in $\operatorname{span}(v_1,\ldots,v_i)$; so by 5.39 both D_w and D_z have upper-triangular matrices with respect to $v_1,\ldots,v_{10}$.

Problem 5E.9. Suppose V is a finite-dimensional nonzero complex vector space. Suppose that $\mathcal{E} \subseteq \mathcal{L}(V)$ is such that S and T commute for all $S,T \in \mathcal{E}$.

- (a) Prove that there is a vector in V that is an eigenvector for every element of $\mathcal{E}$.
- (b) Prove that there is a basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

[This exercise extends 5.78 and 5.80, which consider the case in which $\mathcal{E}$ contains only two elements. For this exercise, $\mathcal{E}$ may contain any number of elements, and $\mathcal{E}$ may even be an infinite set.]

Solution. (a) Induct on $n = \dim V \ge 1$. For $n = 1$, any nonzero v spans V, so $Tv \in \operatorname{span}(v)$ makes v an eigenvector of every $T \in \mathcal{E}$. Let $n > 1$.
(i) If every element of $\mathcal{E}$ is a scalar multiple of I, any nonzero $v \in V$ serves.
(ii) Otherwise choose $S \in \mathcal{E}$ that is not a scalar multiple of I. Since V is a nonzero finite-dimensional complex vector space, S has an eigenvalue λ (5.19); put $U = E(\lambda, S)$, so $U \ne \{0\}$, and $U \ne V$ since $U = V$ would mean $S = \lambda I$. Hence $1 \le \dim U \le n-1$. Every $T \in \mathcal{E}$ commutes with S, so U is invariant under T (5.75), and the restrictions again commute pairwise, since

$$(T_1|_U)(T_2|_U)u = T_1T_2u = T_2T_1u = (T_2|_U)(T_1|_U)u \qquad (u \in U).$$

By the induction hypothesis some nonzero $v \in U$ is an eigenvector of every $T|_U$, hence of every $T \in \mathcal{E}$, as $Tv = (T|_U)v$. (The eigenvalues depend on T; the vector does not.)
(b) Induct on $n = \dim V \ge 1$; for $n = 1$ every 1-by-1 matrix is upper triangular. Let $n > 1$. By (a) there is a nonzero v_1 that is an eigenvector of every element of $\mathcal{E}$, so $U = \operatorname{span}(v_1)$ is invariant under every element of $\mathcal{E}$ and the quotient operators T/U, $(T/U)(v+U) = Tv + U$ (Exercise 38 in Section 5A), are defined on V/U, where $\dim V/U = n - 1 \ge 1$ by 3.105. They commute pairwise:

$$(T_1/U)(T_2/U)(v+U) = T_1T_2v + U = T_2T_1v + U = (T_2/U)(T_1/U)(v+U).$$

So the induction hypothesis gives a basis $u_2 + U, \ldots, u_n + U$ of V/U with respect to which every T/U is upper triangular.
Then $v_1, u_2, \ldots, u_n$ is a basis of V: it has length n, and if $a_1v_1 + a_2u_2 + \cdots + a_nu_n = 0$, applying the quotient map (which sends v_1 to $0 + U$) forces $a_2 = \cdots = a_n = 0$ by independence in V/U, whence $a_1v_1 = 0$ and $a_1 = 0$; now 2.38 applies.
Fix $T \in \mathcal{E}$. Then $Tv_1 \in \operatorname{span}(v_1)$, and for $k \geq 2$ upper-triangularity of $\mathcal{M}(T/U)$ gives, by 5.39, scalars $a_2, \ldots, a_k$ with

$$Tu_k + U = \sum_{j=2}^{k} a_j(u_j + U) = \Big(\sum_{j=2}^{k} a_ju_j\Big) + U,$$

so $Tu_k - \sum_{j=2}^{k} a_ju_j \in U = \operatorname{span}(v_1)$ and $Tu_k \in \operatorname{span}(v_1, u_2, \ldots, u_k)$. By 5.39 again, every $T \in \mathcal{E}$ has an upper-triangular matrix with respect to $v_1, u_2, \ldots, u_n$. Both inductions are on $\dim V$, so $\mathcal{E}$ may be infinite.

Problem 5E.10. Give an example of two commuting operators S, T on a finite-dimensional real vector space such that $S + T$ has an eigenvalue that does not equal an eigenvalue of S plus an eigenvalue of T and ST has an eigenvalue that does not equal an eigenvalue of S times an eigenvalue of T.
[This exercise shows that 5.81 does not hold on real vector spaces.]

Solution. Take $V = \mathbf{R}^2$ and define $S, T \in \mathcal{L}(\mathbf{R}^2)$ by

$$S(x, y) = (-y, x), \qquad T(x, y) = (y, -x).$$

Thus $T = -S$, and S, T commute: $ST = -S^2 = TS$.
Neither has an eigenvalue: $S(x, y) = \lambda(x, y)$ says $-y = \lambda x$ and $x = \lambda y$, so $(\lambda^2 + 1)y = 0$, forcing $y = 0$ (as $\lambda \in \mathbf{R}$) and then $x = \lambda y = 0$; and $T(x, y) = \lambda(x, y)$ is the equation $S(x, y) = (-\lambda)(x, y)$.
But $S + T = 0$ has eigenvalue 0, and $ST = I$ has eigenvalue 1, since

$$(ST)(x, y) = S(y, -x) = (x, y).$$

Since S and T have no eigenvalues at all, no number whatsoever is an eigenvalue of S plus an eigenvalue of T, or an eigenvalue of S times an eigenvalue of T; in particular 0 and 1 are not.

6 Inner Product Spaces

Contents

6.1 Exercises 6A

Problem 6A.1. Prove or give a counterexample: If $v_1, \ldots, v_m \in V$, then

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j, v_k\rangle \geq 0.$$

Solution. True: the double sum equals $\|v_1 + \cdots + v_m\|^2$.
Put $v = v_1 + \cdots + v_m$. Additivity in the second slot (6.6(d)) gives $\sum_k \langle v_j, v_k\rangle = \langle v_j, v\rangle$, and additivity in the first slot (6.2) then gives

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle v_j, v_k\rangle = \sum_{j=1}^{m}\langle v_j, v\rangle = \langle v, v\rangle = \|v\|^2 \geq 0,$$

the norm by 6.7 and the inequality by positivity in 6.2, which also makes the sum a nonnegative real number when $\mathbf{F} = \mathbf{C}$.

Problem 6A.2. Suppose $S \in \mathcal{L}(V)$. Define $\langle \cdot, \cdot\rangle_1$ by

$$\langle u, v\rangle_1 = \langle Su, Sv\rangle$$

for all $u, v \in V$. Show that $\langle \cdot, \cdot\rangle_1$ is an inner product on V if and only if S is injective.

Solution. Of the five properties in 6.2, all but definiteness hold for $\langle \cdot, \cdot\rangle_1$ regardless of S: positivity is $\langle v, v\rangle_1 = \langle Sv, Sv\rangle \geq 0$, conjugate symmetry is inherited from $\langle \cdot, \cdot\rangle$, and linearity of S turns additivity and homogeneity of $\langle \cdot, \cdot\rangle$ in the first slot into the same for $\langle \cdot, \cdot\rangle_1$, as in

$$\begin{aligned}\langle \lambda u + w, v\rangle_1 &= \langle \lambda Su + Sw, Sv\rangle \\ &= \lambda\langle u, v\rangle_1 + \langle w, v\rangle_1.\end{aligned}$$

So $\langle \cdot, \cdot\rangle_1$ is an inner product if and only if it is definite. By definiteness of $\langle \cdot, \cdot\rangle$,

$$\langle v, v\rangle_1 = 0 \iff \langle Sv, Sv\rangle = 0 \iff Sv = 0,$$

so $\{v \in V : \langle v, v\rangle_1 = 0\} = \operatorname{null} S$. Hence $\langle \cdot, \cdot\rangle_1$ is definite if and only if $\operatorname{null} S = \{0\}$, which by 3.15 holds if and only if S is injective.

Problem 6A.3. (a) Show that the function taking an ordered pair $((x_1, x_2), (y_1, y_2))$ of elements of $\mathbf{R}^2$ to $|x_1 y_1| + |x_2 y_2|$ is not an inner product on $\mathbf{R}^2$.
(b) Show that the function taking an ordered pair $((x_1, x_2, x_3), (y_1, y_2, y_3))$ of elements of $\mathbf{R}^3$ to $x_1 y_1 + x_3 y_3$ is not an inner product on $\mathbf{R}^3$.

Solution. (a) Homogeneity in the first slot fails at $\lambda = -1$, $u = v = (1, 0)$: writing φ for the function,

$$\varphi(-u, v) = |(-1)(1)| + 0 = 1 \neq -1 = -\varphi(u, v).$$

So φ is not an inner product on $\mathbf{R}^2$.
(b) Definiteness fails at $v = (0, 1, 0) \neq 0$: writing ψ for the function,

$$\psi(v, v) = 0 \cdot 0 + 0 \cdot 0 = 0.$$

So ψ is not an inner product on $\mathbf{R}^3$.
Method (2) for (b): $\psi(u, v) = \langle Su, Sv\rangle$ for the Euclidean inner product and $S \in \mathcal{L}(\mathbf{R}^3)$ given by $S(x_1, x_2, x_3) = (x_1, 0, x_3)$, which is not injective; apply Exercise 6A.2.

Problem 6A.4. Suppose $T \in \mathcal{L}(V)$ is such that $\|Tv\| \le \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}\, I$ is injective.

Solution. By 3.15 it suffices to show $\operatorname{null}(T - \sqrt{2}\, I) = \{0\}$. If $(T - \sqrt{2}\, I)v = 0$, then $Tv = \sqrt{2}\, v$, so by 6.9(b) and the hypothesis

$$\sqrt{2}\, \|v\| = \|\sqrt{2}\, v\| = \|Tv\| \le \|v\|,$$

whence $(\sqrt{2} - 1)\|v\| \le 0$. Since $\sqrt{2} - 1 > 0$ and norms are nonnegative, $\|v\| = 0$ and thus $v = 0$ by 6.9(a).

Problem 6A.5. Suppose V is a real inner product space.
(a) Show that $\langle u + v, u - v\rangle = \|u\|^2 - \|v\|^2$ for every $u, v \in V$.
(b) Show that if $u, v \in V$ have the same norm, then $u + v$ is orthogonal to $u - v$.
(c) Use (b) to show that the diagonals of a rhombus are perpendicular to each other.

Solution. (a) Expanding by additivity in each slot (6.2 and 6.6(d), with 6.6(e) at $\lambda = -1$),

$$\begin{aligned}\langle u + v, u - v\rangle = \langle u, u\rangle - \langle u, v\rangle& \\ + \langle v, u\rangle - \langle v, v\rangle&,\end{aligned}$$

and the middle terms cancel because $\mathbf{F} = \mathbf{R}$ makes conjugate symmetry plain symmetry. So $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$ by 6.7.
(b) If $\|u\| = \|v\|$, then $\langle u + v, u - v\rangle = \|u\|^2 - \|v\|^2 = 0$ by (a), which is orthogonality by 6.10.
(c) Place the rhombus in $\mathbf{R}^2$ with a vertex at the origin, and let u, v be the sides emanating from it; being a parallelogram, its vertices are $0, u, u + v, v$, its four sides are $\pm u, \pm v$ with lengths $\|u\|, \|v\|, \|u\|, \|v\|$ by 6.9(b), and u, v are linearly independent. Equal side lengths thus say exactly $\|u\| = \|v\|$, while the diagonals are the vectors $u + v$ and $u - v$, both nonzero by independence. By (b) they are orthogonal, hence perpendicular (Exercise 6A.15).

Problem 6A.6. Suppose $u, v \in V$. Prove that $\langle u, v\rangle = 0 \iff \|u\| \le \|u + av\|$ for all $a \in \mathbf{F}$.

Solution. $\Rightarrow$ If $\langle u, v\rangle = 0$ and $a \in \mathbf{F}$, then $\langle u, av\rangle = \overline{a}\langle u, v\rangle = 0$ by 6.6(e), so $u \perp av$ and the Pythagorean theorem 6.12 gives

$$\|u + av\|^2 = \|u\|^2 + \|av\|^2 \ge \|u\|^2.$$

$\Leftarrow$ Suppose $\|u\| \le \|u + av\|$ for all $a \in \mathbf{F}$. If $v = 0$ then $\langle u, v\rangle = 0$ by 6.6(c). Otherwise the orthogonal decomposition 6.13 gives $u = cv + w$ with

$$c = \frac{\langle u, v\rangle}{\|v\|^2}, \quad w = u - cv, \quad \langle w, v\rangle = 0,$$

and then $\langle cv, w\rangle = c\overline{\langle w, v\rangle} = 0$, so 6.12 and 6.9(b) give

$$\|u\|^2 = \|cv\|^2 + \|w\|^2 = |c|^2\|v\|^2 + \|w\|^2.$$

The hypothesis at $a = -c$ reads $\|u\| \le \|u - cv\| = \|w\|$; squaring and substituting yields $|c|^2\|v\|^2 \le 0$. Since $\|v\| \ne 0$ by 6.9(a), we get $c = 0$ and hence $\langle u, v\rangle = 0$.

Problem 6A.7. Suppose $u, v \in V$. Prove that $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbf{R}$ if and only if $\|u\| = \|v\|$.

Solution. Everything follows from the identity

$$\|au + bv\|^2 - \|bu + av\|^2 = (a^2 - b^2)\big(\|u\|^2 - \|v\|^2\big) \qquad (a, b \in \mathbf{R}).$$

To prove it, expand by additivity in both slots (6.2, 6.6(d)) and use $\overline{a} = a$, $\overline{b} = b$ together with 6.6(e), 6.9(b) and conjugate symmetry, so that the cross terms $ab\langle u, v\rangle + ab\overline{\langle u, v\rangle}$ combine:

$$\|au + bv\|^2 = a^2\|u\|^2 + b^2\|v\|^2 + 2ab\,\mathrm{Re}\langle u, v\rangle.$$

Swapping a and b fixes the cross term and exchanges the first two, giving the identity.
If $\|u\| = \|v\|$, the right side vanishes for all real a, b, so $\|au + bv\| = \|bu + av\|$. Conversely, $a = 1, b = 0$ gives $\|u\| = \|v\|$ directly.

Problem 6A.8. Suppose $a, b, c, x, y \in \mathbb{R}$ and $a^2 + b^2 + c^2 + x^2 + y^2 \leq 1$. Prove that

$$a + b + c + 4x + 9y \leq 10.$$

Solution. Apply Cauchy–Schwarz (6.14) in $\mathbb{R}^5$ with its Euclidean inner product to

$$u = (a, b, c, x, y), \qquad w = (1, 1, 1, 4, 9).$$

Here $\|w\| = \sqrt{1 + 1 + 1 + 16 + 81} = 10$, while the hypothesis says $\|u\|^2 \leq 1$, so $\|u\| \leq 1$. Hence

$$a + b + c + 4x + 9y = \langle u, w\rangle \leq \|u\|\,\|w\| \leq 10.$$

Problem 6A.9. Suppose $u, v \in V$ and $\|u\| = \|v\| = 1$ and $\langle u, v\rangle = 1$. Prove that $u = v$.

Solution. Conjugate symmetry gives $\langle v, u\rangle = \overline{\langle u, v\rangle} = 1$, so expanding $\|u - v\|^2$ by 6.7 and 6.6(a), (d), (e),

$$\begin{aligned}
\|u - v\|^2 &= \langle u - v, u - v\rangle \\
&= \langle u, u\rangle - \langle u, v\rangle - \langle v, u\rangle + \langle v, v\rangle \\
&= \|u\|^2 - \langle u, v\rangle - \langle v, u\rangle + \|v\|^2 \\
&= 1 - 1 - 1 + 1 \\
&= 0.
\end{aligned}$$

By 6.9(a), $\|u - v\| = 0$ implies $u - v = 0$, that is, $u = v$.

Problem 6A.10. Suppose $u, v \in V$ and $\|u\| \leq 1$ and $\|v\| \leq 1$. Prove that

$$\sqrt{1 - \|u\|^2}\,\sqrt{1 - \|v\|^2} \leq 1 - |\langle u, v\rangle|.$$

Solution. Write $a = \|u\| \in [0, 1]$ and $b = \|v\| \in [0, 1]$. Cauchy–Schwarz (6.14) gives $|\langle u, v\rangle| \leq ab$, so $1 - ab \leq 1 - |\langle u, v\rangle|$ and it suffices to prove the numerical inequality

$$\sqrt{1 - a^2}\,\sqrt{1 - b^2} + ab \leq 1.$$

Apply 6.14 again, in $\mathbb{R}^2$ with the Euclidean inner product, to

$$s = \big(\sqrt{1 - a^2},\, a\big), \quad t = \big(\sqrt{1 - b^2},\, b\big),$$

both of which are unit vectors, since $\|s\| = \sqrt{(1 - a^2) + a^2} = 1$ and likewise $\|t\| = 1$. Hence

$$\sqrt{1 - a^2}\,\sqrt{1 - b^2} + ab = \langle s, t\rangle \leq \|s\|\,\|t\| = 1.$$

Method (2) for the numerical step: both sides of $\sqrt{1 - a^2}\sqrt{1 - b^2} \leq 1 - ab$ are nonnegative since $ab \leq 1$, so squaring is equivalent, and $(1 - a^2)(1 - b^2) \leq (1 - ab)^2$ reduces to $0 \leq (a - b)^2$.

Problem 6A.11. Find vectors $u, v \in \mathbb{R}^2$ such that u is a scalar multiple of $(1,3)$, v is orthogonal to $(1,3)$, and $(1,2) = u + v$.

Solution. Take
$$u = \left(\tfrac{7}{10}, \tfrac{21}{10}\right), \qquad v = \left(\tfrac{3}{10}, -\tfrac{1}{10}\right).$$
These come from the orthogonal decomposition 6.13 with $c = \langle (1,2), (1,3)\rangle / \|(1,3)\|^2 = 7/10$, so that $u = c(1,3)$ and $v = (1,2) - u$. Then u is a multiple of $(1,3)$, $\langle v, (1,3)\rangle = \frac{3}{10} - \frac{3}{10} = 0$, and $u + v = (1,2)$. (Check!)

Problem 6A.12. Suppose a, b, c, d are positive numbers.
(a) Prove that $(a+b+c+d)\left(\dfrac{1}{a} + \dfrac{1}{b} + \dfrac{1}{c} + \dfrac{1}{d}\right) \geq 16$.
(b) For which positive numbers a, b, c, d is the inequality above an equality?

Solution. Apply Cauchy–Schwarz (6.14) in $\mathbb{R}^4$ to
$$s = \left(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}\right),$$
$$t = \left(\tfrac{1}{\sqrt{a}}, \tfrac{1}{\sqrt{b}}, \tfrac{1}{\sqrt{c}}, \tfrac{1}{\sqrt{d}}\right),$$
which are defined because $a, b, c, d > 0$, and satisfy $\langle s, t\rangle = 4$, $\|s\|^2 = a+b+c+d$, and $\|t\|^2 = \frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}$.
(a) Squaring $|\langle s, t\rangle| \leq \|s\|\,\|t\|$ gives
$$16 = \langle s, t\rangle^2 \leq \|s\|^2\,\|t\|^2 = (a+b+c+d)\left(\tfrac{1}{a} + \tfrac{1}{b} + \tfrac{1}{c} + \tfrac{1}{d}\right).$$
(b) Exactly when $a = b = c = d$. If so, the product is $(4a)(4/a) = 16$. Conversely, equality forces equality in 6.14, so by its equality condition one of s, t is a scalar multiple of the other; both are nonzero, so $t = \lambda s$ for some $\lambda \in \mathbb{R}$. Comparing coordinates,
$$\lambda = \tfrac{1}{a} = \tfrac{1}{b} = \tfrac{1}{c} = \tfrac{1}{d},$$
and positivity gives $a = b = c = d$.

Problem 6A.13. Show that the square of an average is less than or equal to the average of the squares. More precisely, show that if $a_1, \ldots, a_n \in \mathbb{R}$, then the square of the average of $a_1, \ldots, a_n$ is less than or equal to the average of $a_1^2, \ldots, a_n^2$.

Solution. Apply Cauchy–Schwarz (6.14) in $\mathbb{R}^n$ to $a = (a_1, \ldots, a_n)$ and $e = (1, \ldots, 1)$, for which $\langle a, e\rangle = a_1 + \cdots + a_n$, $\|a\|^2 = a_1^2 + \cdots + a_n^2$, and $\|e\|^2 = n$. Squaring $|\langle a, e\rangle| \leq \|a\|\,\|e\|$ gives
$$(a_1 + \cdots + a_n)^2 \leq n\,(a_1^2 + \cdots + a_n^2),$$
and dividing by $n^2 > 0$ gives
$$\left(\frac{a_1 + \cdots + a_n}{n}\right)^2 \leq \frac{a_1^2 + \cdots + a_n^2}{n},$$
the square of the average bounded by the average of the squares.

Problem 6A.14. Suppose $v \in V$ and $v \neq 0$. Prove that $v/\|v\|$ is the unique closest element on the unit sphere of V to v. More precisely, prove that if $u \in V$ and $\|u\| = 1$, then
$$\left\| v - \frac{v}{\|v\|} \right\| \leq \|v - u\|,$$

with equality only if $u = v/\|v\|$.

Solution. The left side equals $\big|\,\|v\| - 1\,\big|$: since $\|v\| > 0$ by 6.9(a), 6.9(b) gives

$$\left\| v - \frac{v}{\|v\|} \right\| = \left| 1 - \frac{1}{\|v\|} \right| \|v\| = \big|\,\|v\| - 1\,\big|.$$

Now let $\|u\| = 1$. Expanding by 6.7 and using conjugate symmetry to write the cross terms as $2\operatorname{Re}\langle v, u\rangle$,

$$\|v - u\|^2 = \|v\|^2 - 2\operatorname{Re}\langle v, u\rangle + 1,$$

while $\operatorname{Re}\langle v, u\rangle \le |\langle v, u\rangle| \le \|v\|\,\|u\| = \|v\|$ by Cauchy–Schwarz (6.14). Hence

$$\|v - u\|^2 \ge \|v\|^2 - 2\|v\| + 1 = \big(\|v\| - 1\big)^2,$$

and taking nonnegative square roots gives the asserted inequality.
Equality forces $\operatorname{Re}\langle v, u\rangle = \|v\| = \|v\|\,\|u\|$, so both inequalities above are equalities. In particular Cauchy–Schwarz is an equality, so by the equality condition in 6.14 one of v, u is a multiple of the other; as both are nonzero, $v = \lambda u$ with $\lambda \ne 0$. Then $\langle v, u\rangle = \lambda\|u\|^2 = \lambda$ and $\|v\| = |\lambda|$, so $\operatorname{Re}\lambda = |\lambda|$ forces $\lambda > 0$, whence $\lambda = \|v\|$ and

$$u = \frac{v}{\lambda} = \frac{v}{\|v\|}.$$

Problem 6A.15. Suppose u, v are nonzero vectors in $\mathbf{R}^2$. Prove that

$$\langle u, v\rangle = \|u\|\,\|v\| \cos\theta,$$

where θ is the angle between u and v (thinking of u and v as arrows with initial point at the origin). Hint: Use the law of cosines on the triangle formed by u, v, and $u - v$.

Solution. Expanding in the (symmetric, since real) Euclidean inner product gives the identity

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle,$$

in which $\|u\|, \|v\|, \|u - v\|$ are the ordinary Euclidean lengths. Let $\theta \in [0, \pi]$ be the angle between u and v.
(i) u, v not scalar multiples of each other. Then $0, u, v$ form a genuine triangle, with sides $\|u\|, \|v\|$ at the vertex 0, included angle θ, and opposite side $\|u - v\|$, so the law of cosines gives

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\,\|v\| \cos\theta.$$

Comparing with the identity yields $\langle u, v\rangle = \|u\|\,\|v\| \cos\theta$.
(ii) $v = tu$ with $t \ne 0$ (both vectors being nonzero). Then $\|v\| = |t|\,\|u\|$ by 6.9(b) and $\cos\theta = \operatorname{sign}(t)$, since $\theta = 0$ for $t > 0$ and $\theta = \pi$ for $t < 0$; hence

$$\langle u, v\rangle = t\|u\|^2 = \operatorname{sign}(t)\,\|u\|\,|t|\,\|u\| = \|u\|\,\|v\| \cos\theta.$$

Problem 6A.16. The angle between two vectors (thought of as arrows with initial point at the origin) in $\mathbf{R}^2$ or $\mathbf{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbf{R}^n$ for $n > 3$. Thus the angle between two nonzero vectors $x, y \in \mathbf{R}^n$ is defined to be

$$\arccos \frac{\langle x, y\rangle}{\|x\|\,\|y\|},$$

where the motivation for this definition comes from Exercise 15. Explain why the Cauchy–Schwarz inequality is needed to show that this definition makes sense.

Solution. Because arccos has domain exactly $[-1,1]$, being the inverse of cos restricted to $[0,\pi]$, the definition means nothing unless

$$\frac{\langle x,y\rangle}{\|x\|\,\|y\|}\in[-1,1].$$

The denominator is positive because $x,y\neq 0$ gives $\|x\|,\|y\|\neq 0$ by 6.9(a); and it is exactly the Cauchy–Schwarz inequality (6.14), $|\langle x,y\rangle|\leq\|x\|\,\|y\|$, that puts the quotient in $[-1,1]$ after dividing by $\|x\|\,\|y\|>0$. Nothing in the definitions of $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ alone gives that bound, so without 6.14 the arccosine could be undefined.

Problem 6A.17. Prove that

$$\left(\sum_{k=1}^{n}a_kb_k\right)^2\leq\left(\sum_{k=1}^{n}ka_k^2\right)\left(\sum_{k=1}^{n}\frac{b_k^2}{k}\right)$$

for all real numbers $a_1,\dots,a_n$ and $b_1,\dots,b_n$.

Solution. Split $a_kb_k=(\sqrt{k}\,a_k)(b_k/\sqrt{k})$, legitimate since $k\geq 1>0$, and apply Cauchy–Schwarz in coordinate form 6.16(a) to

$$x=\left(\sqrt{k}\,a_k\right)_{k=1}^{n},\qquad y=\left(b_k/\sqrt{k}\right)_{k=1}^{n},$$

for which $x_ky_k=a_kb_k$, $x_k^2=ka_k^2$, and $y_k^2=b_k^2/k$. Substituting into $(x_1y_1+\cdots+x_ny_n)^2\leq(x_1^2+\cdots+x_n^2)(y_1^2+\cdots+y_n^2)$ gives

$$\left(\sum_{k=1}^{n}a_kb_k\right)^2\leq\left(\sum_{k=1}^{n}ka_k^2\right)\left(\sum_{k=1}^{n}\frac{b_k^2}{k}\right).$$

Problem 6A.18. (a) Suppose $f\colon[1,\infty)\to[0,\infty)$ is continuous. Show that

$$\left(\int_1^\infty f\right)^2\leq\int_1^\infty x^2\big(f(x)\big)^2\,dx.$$

(b) For which continuous functions $f\colon[1,\infty)\to[0,\infty)$ is the inequality in (a) an equality with both sides finite?

Solution. Write $F(b)=\int_1^b f$ and $G(b)=\int_1^b x^2f(x)^2\,dx$; both integrands are continuous and nonnegative, so F,G are nondecreasing and the improper integrals $L=\lim_{b\to\infty}F(b)$ and $M=\lim_{b\to\infty}G(b)$ exist in $[0,\infty]$.
(a) Apply Cauchy–Schwarz in the integral form 6.16(b) on $[1,b]$ to $g(x)=1/x$ and $h(x)=xf(x)$, both continuous there, with $gh=f$:

$$F(b)^2\leq\left(\int_1^b\frac{dx}{x^2}\right)G(b)=\left(1-\tfrac{1}{b}\right)G(b)\leq G(b)\leq M.$$

Letting $b\to\infty$ gives $L^2\leq M$, which is the claim (trivially so if $M=\infty$).
(b) Exactly the functions $f(x)=c/x^2$ with $c\geq 0$ constant. For such f,

$$\int_1^\infty f=c,\qquad\int_1^\infty x^2f(x)^2\,dx=c^2\int_1^\infty\frac{dx}{x^2}=c^2,$$

both finite and equal after squaring the first. Conversely suppose $L^2=M<\infty$, and set $c=L\geq 0$. Expanding the square in $\Phi(b)=\int_1^b\left(xf(x)-\frac{c}{x}\right)^2dx$, using $xf(x)\cdot\frac{c}{x}=cf(x)$,

$$\Phi(b)=G(b)-2cF(b)+c^2\left(1-\tfrac{1}{b}\right)\longrightarrow M-2L^2+L^2=0.$$

Since Φ is nondecreasing and nonnegative with limit 0, it vanishes identically. A continuous nonnegative function with zero integral over $[1, b]$ is identically 0 there, so $xf(x) = c/x$, that is, $f(x) = c/x^2$, for all $x \geq 1$.

Problem 6A.19. Suppose $v_1, \ldots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda|^2 \leq \sum_{j=1}^{n} \sum_{k=1}^{n} \left|\mathcal{M}(T)_{j,k}\right|^2,$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \ldots, v_n$.

Solution. Write $A_{j,k} = \mathcal{M}(T)_{j,k}$, so $Tv_k = \sum_j A_{j,k} v_j$ by 3.31, and let $v = \sum_k c_k v_k$ be an eigenvector for λ, so that $\sum_k |c_k|^2 > 0$ since $v \neq 0$. Expanding $Tv = \lambda v$ in the basis and comparing coefficients (the representation being unique) gives

$$\sum_{k=1}^{n} A_{j,k} c_k = \lambda c_j \qquad \text{for } j = 1, \ldots, n.$$

Fix j and apply Cauchy–Schwarz (6.14) in $\mathbf{F}^n$ with $\langle x, y \rangle = \sum_k x_k \overline{y_k}$ to $x = (c_1, \ldots, c_n)$ and $y = \left(\overline{A_{j,1}}, \ldots, \overline{A_{j,n}}\right)$, for which $\langle x, y \rangle = \sum_k A_{j,k} c_k$ and $\|y\|^2 = \sum_k |A_{j,k}|^2$:

$$|\lambda|^2 |c_j|^2 \leq \left(\sum_{k=1}^{n} |A_{j,k}|^2 \right) \left(\sum_{k=1}^{n} |c_k|^2 \right).$$

Summing over j and dividing by $\sum_k |c_k|^2 > 0$ gives $|\lambda|^2 \leq \sum_{j,k} |A_{j,k}|^2$.

Problem 6A.20. Prove that if $u, v \in V$, then $\big|\|u\| - \|v\|\big| \leq \|u - v\|$.
The inequality above is called the reverse triangle inequality. For the reverse triangle inequality when $V = \mathbf{C}$, see Exercise 2 in Chapter 4.

Solution. It suffices to bound $\pm(\|u\| - \|v\|)$ by $\|u - v\|$. The triangle inequality (6.17) applied to $u = (u - v) + v$ gives

$$\|u\| \leq \|u - v\| + \|v\|,$$

so $\|u\| - \|v\| \leq \|u - v\|$; interchanging u and v gives $\|v\| - \|u\| \leq \|v - u\| = \|u - v\|$, the last equality by 6.9(b) with $\lambda = -1$. Hence $\big|\|u\| - \|v\|\big| \leq \|u - v\|$.

Problem 6A.21. Suppose $u, v \in V$ are such that

$$\|u\| = 3, \qquad \|u + v\| = 4, \qquad \|u - v\| = 6.$$

What number does $\|v\|$ equal?

Solution. $\|v\| = \sqrt{17}$. The parallelogram equality (6.21) gives

$$4^2 + 6^2 = 2\left(3^2 + \|v\|^2\right),$$

so $2\|v\|^2 = 52 - 18 = 34$ and $\|v\|^2 = 17$.

Problem 6A.22. Show that if $u, v \in V$, then

$$\|u+v\|\,\|u-v\| \le \|u\|^2 + \|v\|^2.$$

Solution. Combine $ab \le \frac{a^2+b^2}{2}$ (from $0 \le (a-b)^2$) at $a = \|u+v\|$, $b = \|u-v\|$ with the parallelogram equality (6.21):

$$\|u+v\|\,\|u-v\| \le \frac{\|u+v\|^2 + \|u-v\|^2}{2} = \|u\|^2 + \|v\|^2.$$

Problem 6A.23. Suppose $v_1, \ldots, v_m \in V$ are such that $\|v_k\| \le 1$ for each $k = 1, \ldots, m$. Show that there exist $a_1, \ldots, a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m\| \le \sqrt{m}.$$

Solution. Average over all 2^m sign vectors. Put $A = \{1,-1\}^m$ and $N(a) = \|a_1 v_1 + \cdots + a_m v_m\|^2$ for $a \in A$. Since each a_k is real, expanding by 6.6 in each slot gives

$$N(a) = \sum_{j=1}^{m} \sum_{k=1}^{m} a_j a_k \langle v_j, v_k \rangle,$$

so summing over $a \in A$ and interchanging the order of summation,

$$\sum_{a \in A} N(a) = \sum_{j=1}^{m} \sum_{k=1}^{m} \Big(\sum_{a \in A} a_j a_k \Big) \langle v_j, v_k \rangle.$$

The inner sum is 2^m when $j = k$ (each term is $a_j^2 = 1$) and 0 when $j \neq k$, because flipping the j^{th} coordinate is a fixed-point-free involution of A that negates $a_j a_k$. Hence

$$\sum_{a \in A} N(a) = 2^m \sum_{k=1}^{m} \|v_k\|^2 \le 2^m m.$$

So some $N(a) \le m$, since otherwise the sum of the 2^m terms would exceed $2^m m$; for that a, $\|a_1 v_1 + \cdots + a_m v_m\| \le \sqrt{m}$.

Problem 6A.24. Prove or give a counterexample: If $\|\cdot\|$ is the norm associated with an inner product on $\mathbb{R}^2$, then there exists $(x,y) \in \mathbb{R}^2$ such that $\|(x,y)\| \neq \max\{|x|, |y|\}$.

Solution. True: no inner product on $\mathbb{R}^2$ has $\max\{|x|,|y|\}$ as its norm, since that function violates the parallelogram equality (6.21) at $u = (1,0)$, $v = (0,1)$. Indeed $u + v = (1,1)$ and $u - v = (1,-1)$, so all four of $\|u\|, \|v\|, \|u+v\|, \|u-v\|$ equal 1 and

$$\|u+v\|^2 + \|u-v\|^2 = 2 \neq 4 = 2\big(\|u\|^2 + \|v\|^2\big).$$

Hence for any inner product norm on $\mathbb{R}^2$, at least one of these four vectors has norm different from the maximum of its coordinates absolute values.

Problem 6A.25. Suppose $p > 0$. Prove that there is an inner product on $\mathbb{R}^2$ such that the associated norm is given by

$$\|(x,y)\| = \big(|x|^p + |y|^p\big)^{1/p}$$

for all $(x,y) \in \mathbb{R}^2$ if and only if $p = 2$.

Solution. Write $N_p(x,y) = (|x|^p + |y|^p)^{1/p}$.
$\Leftarrow$ For $p=2$, the Euclidean inner product on $\mathbb{R}^2$ (an inner product by 6.3(a)) has associated norm $\sqrt{x^2+y^2} = N_2(x,y)$ by 6.7.
$\Rightarrow$ If N_p comes from an inner product, the parallelogram equality (6.21) applies to $u=(1,0)$ and $v=(0,1)$, for which $\|u\| = \|v\| = 1$ and $\|u+v\| = \|u-v\| = 2^{1/p}$:

$$2 \cdot 2^{2/p} = 2^{1+2/p} = 2(1+1) = 2^2.$$

Injectivity of $t \mapsto 2^t$ gives $1 + 2/p = 2$, so $p = 2$.

Problem 6A.26. Suppose V is a real inner product space. Prove that

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$

for all $u, v \in V$.

Solution. Expanding by additivity in both slots (6.2, 6.6(d), 6.6(e)) and using $\langle v, u\rangle = \langle u, v\rangle$, which is conjugate symmetry over $\mathbf{R}$,

$$\|u \pm v\|^2 = \|u\|^2 + \|v\|^2 \pm 2\langle u, v\rangle.$$

Subtracting cancels $\|u\|^2 + \|v\|^2$ and leaves $\|u+v\|^2 - \|u-v\|^2 = 4\langle u, v\rangle$.

Problem 6A.27. Suppose V is a complex inner product space. Prove that

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4}$$

for all $u, v \in V$.

Solution. Everything follows from the expansion

$$\|u+w\|^2 = \|u\|^2 + \|w\|^2 + \langle u, w\rangle + \overline{\langle u, w\rangle} = \|u\|^2 + \|w\|^2 + 2\,\mathrm{Re}\langle u, w\rangle,$$

valid for all $u, w \in V$ by conjugate symmetry and $z + \overline{z} = 2\,\mathrm{Re}\, z$. Taking $w = v$ and $w = -v$ (using $\|-v\| = \|v\|$ and $\langle u, -v\rangle = -\langle u, v\rangle$, by 6.6(e)) gives

$$\|u+v\|^2 - \|u-v\|^2 = 4\,\mathrm{Re}\langle u, v\rangle.$$

Taking $w = iv$ and $w = -iv$ instead, and using $\|\pm iv\| = |\pm i|\,\|v\| = \|v\|$ together with

$$\begin{aligned} \langle u, \pm iv\rangle &= \overline{\pm i}\,\langle u, v\rangle = \mp i\langle u, v\rangle, \\ \mathrm{Re}\big(-i\langle u, v\rangle\big) &= \mathrm{Im}\langle u, v\rangle \end{aligned}$$

(by 6.6(e), and writing $\langle u, v\rangle = a + bi$, so $-i(a+bi) = b - ai$), gives

$$\|u+iv\|^2 - \|u-iv\|^2 = 4\,\mathrm{Im}\langle u, v\rangle.$$

Hence the numerator on the right side of the asserted formula is

$$4\,\mathrm{Re}\langle u, v\rangle + 4\big(\mathrm{Im}\langle u, v\rangle\big)i = 4\langle u, v\rangle,$$

and dividing by 4 gives the identity.

Problem 6A.28. A norm on a vector space U is a function

$$\|\cdot\|\colon U \to [0,\infty)$$

such that $\|u\| = 0$ if and only if $u = 0$, $\|\alpha u\| = |\alpha|\|u\|$ for all $\alpha \in \mathbb{F}$ and all $u \in U$, and $\|u+v\| \le \|u\| + \|v\|$ for all $u, v \in U$. Prove that a norm satisfying the parallelogram equality comes from an inner product (in other words, show that if $\|\cdot\|$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle\cdot,\cdot\rangle$ on U such that $\|u\| = \langle u,u\rangle^{1/2}$ for all $u \in U$).

Solution. Take

$$\varphi(u,v) = \frac{\|u+v\|^2 - \|u-v\|^2}{4},$$

which is the required inner product when $\mathbb{F} = \mathbb{R}$; when $\mathbb{F} = \mathbb{C}$, take $\langle u,v\rangle = \varphi(u,v) + \varphi(u,iv)\,i$, the expression of Exercise 27. Write (P) for the parallelogram equality. Two consequences of the norm axioms are used repeatedly: $\|-u\| = \|u\|$ (homogeneity with $\alpha = -1$), and the reverse triangle inequality

$$\big|\,\|a\| - \|b\|\,\big| \le \|a-b\| \tag{R}$$

(from $\|a\| \le \|a-b\| + \|b\|$ and the same with a, b interchanged).

(i) Symmetry, positivity, definiteness. Since $\|v-u\| = \|u-v\|$ we get $\varphi(v,u) = \varphi(u,v)$, and

$$\varphi(u,u) = \frac{\|2u\|^2 - \|0\|^2}{4} = \|u\|^2 \ge 0,$$

which vanishes exactly when $u = 0$. Also $\varphi(u,0) = \varphi(0,u) = 0$ and $\varphi(-u,v) = -\varphi(u,v)$. (Check!)

(ii) Additivity in the first slot. Apply (P) to the pair $u+w, v+w$ and to the pair $u-w, v-w$; both pairs have difference $u-v$, so subtracting cancels $\|u-v\|^2$ and leaves

$$\begin{aligned}\|u+v+2w\|^2 - \|u+v-2w\|^2 = 2\big(&\|u+w\|^2 - \|u-w\|^2\big)\\ &+ 2\big(\|v+w\|^2 - \|v-w\|^2\big),\end{aligned}$$

that is, $\varphi(u+v,2w) = 2\big(\varphi(u,w) + \varphi(v,w)\big)$. Setting $v = 0$ here gives $\varphi(u,2w) = 2\varphi(u,w)$; applying that with u replaced by $u+v$ and comparing yields

$$\varphi(u+v,w) = \varphi(u,w) + \varphi(v,w), \tag{4}$$

and by symmetry φ is additive in the second slot too.

(iii) Real homogeneity. From (4) by induction, together with $\varphi(0,v) = 0$ and $\varphi(-u,v) = -\varphi(u,v)$, we get $\varphi(kw,v) = k\varphi(w,v)$ for every $k \in \mathbb{Z}$. Taking $w = \frac{1}{n}u$ and $k = n$ gives $\varphi(\frac{1}{n}u,v) = \frac{1}{n}\varphi(u,v)$, so $\varphi(\frac{m}{n}u,v) = \frac{m}{n}\varphi(u,v)$ for all $m \in \mathbb{Z}$ and positive integers n. By (R),

$$\big|\,\|tu \pm v\| - \|su \pm v\|\,\big| \le \|(t-s)u\| = |t-s|\,\|u\|,$$

so $t \mapsto \varphi(tu,v)$ is continuous; agreeing with the continuous function $t \mapsto t\varphi(u,v)$ on the dense set $\mathbb{Q}$, it agrees on all of $\mathbb{R}$:

$$\varphi(tu,v) = t\,\varphi(u,v) \qquad \text{for all } t \in \mathbb{R}. \tag{5}$$

For $\mathbb{F} = \mathbb{R}$ this completes the proof: (i), (ii), (iii) are the five conditions of 6.2, and $\varphi(u,u)^{1/2} = \|u\|$.

(iv) The complex case. Restricting scalars to $\mathbb{R}$ leaves $\|\cdot\|$ a norm satisfying (P), so (i), (ii), (iii) apply verbatim and φ is a symmetric $\mathbb{R}$-bilinear form with $\varphi(u,u) = \|u\|^2$. Complex homogeneity of the norm gives

$$\varphi(iu,iv) = \frac{|i|^2\|u+v\|^2 - |i|^2\|u-v\|^2}{4} = \varphi(u,v), \tag{6}$$

whence, using $i \cdot iu = -u$, then (5) in the second slot, then symmetry,

$$\varphi(v,iu) = \varphi(iv, i\cdot iu) = \varphi(iv,-u) = -\varphi(iv,u) = -\varphi(u,iv). \tag{7}$$

Since $\|u \pm iu\| = |1 \pm i| \, \|u\| = \sqrt{2} \, \|u\|$, we get $\varphi(u, iu) = 0$ and hence

$$\langle u, u \rangle = \varphi(u, u) = \|u\|^2,$$

giving positivity, definiteness, and the required identity $\|u\| = \langle u, u \rangle^{1/2}$. Additivity in the first slot is immediate from (4) applied to both terms, and conjugate symmetry follows from symmetry of φ, (7), and the fact that φ is real valued:

$$\langle v, u \rangle = \varphi(u, v) - \varphi(u, iv) \, i = \overline{\langle u, v \rangle}.$$

Finally (5) gives $\langle tu, v \rangle = t \langle u, v \rangle$ for $t \in \mathbb{R}$, while (6), symmetry and (7) give

$$\langle iu, v \rangle = \varphi(iu, v) + \varphi(u, v) \, i = -\varphi(u, iv) + \varphi(u, v) \, i = i \langle u, v \rangle,$$

so for $\alpha = a + bi$ with $a, b \in \mathbb{R}$, additivity yields

$$\langle \alpha u, v \rangle = a \langle u, v \rangle + b \langle iu, v \rangle = (a + bi) \langle u, v \rangle = \alpha \langle u, v \rangle.$$

All five conditions of 6.2 hold.

Problem 6A.29. Suppose $V_1, \ldots, V_m$ are inner product spaces. Show that the equation

$$\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m) \rangle = \langle u_1, v_1 \rangle + \cdots + \langle u_m, v_m \rangle$$

defines an inner product on $V_1 \times \cdots \times V_m$.
In the expression above on the right, for each $k = 1, \ldots, m$, the inner product $\langle u_k, v_k \rangle$ denotes the inner product on V_k. Each of the spaces $V_1, \ldots, V_m$ may have a different inner product, even though the same notation is used here.

Solution. Each of the five conditions of 6.2 holds coordinatewise. Write $\langle \cdot, \cdot \rangle_k$ for the inner product on V_k (all over the same field $\mathbf{F}$, as is implicit in forming the product), so that for $u = (u_1, \ldots, u_m)$ and $v = (v_1, \ldots, v_m)$ in $V = V_1 \times \cdots \times V_m$ the asserted formula reads $\langle u, v \rangle = \sum_{k=1}^{m} \langle u_k, v_k \rangle_k \in \mathbf{F}$.
Positivity and definiteness: $\langle v, v \rangle = \sum_{k=1}^{m} \langle v_k, v_k \rangle_k$ is a finite sum of nonnegative real numbers, hence is nonnegative, and it equals 0 exactly when every term does, which by definiteness in each V_k happens exactly when every $v_k = 0$, that is, when $v = 0$.
Additivity and homogeneity in the first slot: because the operations on V are coordinatewise, $u + v$ has k-th coordinate $u_k + v_k$ and λu has k-th coordinate λu_k, so for $w = (w_1, \ldots, w_m)$ and $\lambda \in \mathbf{F}$,

$$\begin{aligned} \langle u + v, w \rangle &= \sum_{k=1}^{m} \big(\langle u_k, w_k \rangle_k + \langle v_k, w_k \rangle_k \big) = \langle u, w \rangle + \langle v, w \rangle, \\ \langle \lambda u, v \rangle &= \sum_{k=1}^{m} \lambda \langle u_k, v_k \rangle_k = \lambda \langle u, v \rangle. \end{aligned}$$

Conjugate symmetry: since conjugation is additive,

$$\langle u, v \rangle = \sum_{k=1}^{m} \overline{\langle v_k, u_k \rangle_k} = \overline{\sum_{k=1}^{m} \langle v_k, u_k \rangle_k} = \overline{\langle v, u \rangle}.$$

Problem 6A.30. Suppose V is a real inner product space. For $u, v, w, x \in V$, define

$$\langle u + iv, w + ix \rangle_{\mathbf{C}} = \langle u, w \rangle + \langle v, x \rangle + \big(\langle v, w \rangle - \langle u, x \rangle \big) i.$$

(a) Show that $\langle \cdot, \cdot \rangle_{\mathbf{C}}$ makes $V_{\mathbf{C}}$ into a complex inner product space.

(b) Show that if $u, v \in V$, then

$$\langle u, v\rangle_{\mathbf{C}} = \langle u, v\rangle \quad \text{and} \quad \|u + iv\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2.$$

See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$.

Solution. (a) The five conditions of 6.2 all reduce to symmetry and bilinearity of $\langle\cdot,\cdot\rangle$ on the real space V. Throughout, elements of $V_{\mathbf{C}}$ are written uniquely as $u + iv$ with $u, v \in V$ (Exercise 1B.8), with

$$(a + bi)(u + iv) = (au - bv) + i(av + bu), \qquad a, b \in \mathbf{R}.$$

Positivity and definiteness. For $z = u + iv$, the defining formula with $w = u$, $x = v$ gives, since $\langle v, u\rangle = \langle u, v\rangle$ kills the imaginary part,

$$\langle z, z\rangle_{\mathbf{C}} = \langle u, u\rangle + \langle v, v\rangle = \|u\|^2 + \|v\|^2 \geq 0,$$

which equals 0 if and only if $u = v = 0$ (by 6.9(a)), that is, if and only if $z = 0$.
Additivity in the first slot. Since $z_1 + z_2 = (u_1 + u_2) + i(v_1 + v_2)$, applying additivity of $\langle\cdot,\cdot\rangle$ to each of the four terms and regrouping gives $\langle z_1 + z_2, z\rangle_{\mathbf{C}} = \langle z_1, z\rangle_{\mathbf{C}} + \langle z_2, z\rangle_{\mathbf{C}}$. (Check!)
Homogeneity in the first slot. For $a \in \mathbf{R}$ we have $az = (au) + i(av)$, so all four terms of the defining formula scale by a and $\langle az, z'\rangle_{\mathbf{C}} = a\langle z, z'\rangle_{\mathbf{C}}$. For i, note $iz = (-v) + iu$, so with $z' = w + ix$,

$$\begin{aligned}\langle iz, z'\rangle_{\mathbf{C}} &= \big(-\langle v, w\rangle + \langle u, x\rangle\big) + \big(\langle u, w\rangle + \langle v, x\rangle\big)i \\ &= i\langle z, z'\rangle_{\mathbf{C}},\end{aligned}$$

since multiplying $\alpha + \beta i$ by i gives $-\beta + \alpha i$. Hence for $\lambda = a + bi$, writing $\lambda z = az + b(iz)$ and using the additivity just proved, $\langle \lambda z, z'\rangle_{\mathbf{C}} = \lambda\langle z, z'\rangle_{\mathbf{C}}$.
Conjugate symmetry. Conjugating the defining formula for $\langle z', z\rangle_{\mathbf{C}}$ flips the sign of $\langle x, u\rangle - \langle w, v\rangle$, and symmetry of $\langle\cdot,\cdot\rangle$ in all four terms then turns the result into $\langle z, z'\rangle_{\mathbf{C}}$. (Check!)
(b) Identifying $u \in V$ with $u + i0$, the defining formula and $\langle 0, y\rangle = \langle y, 0\rangle = 0$ (6.6(b), 6.6(c)) give

$$\langle u, v\rangle_{\mathbf{C}} = \langle u, v\rangle + \langle 0, 0\rangle + \big(\langle 0, v\rangle - \langle u, 0\rangle\big)i = \langle u, v\rangle,$$

and the positivity computation in (a) is exactly $\|u + iv\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2$.

Problem 6A.31. Suppose $u, v, w \in V$. Prove that

$$\Big\|w - \tfrac{1}{2}(u + v)\Big\|^2 = \frac{\|w - u\|^2 + \|w - v\|^2}{2} - \frac{\|u - v\|^2}{4}.$$

Solution. This is the parallelogram equality 6.21 applied to $a = w - u$ and $b = w - v$, for which

$$a + b = 2\Big(w - \tfrac{1}{2}(u + v)\Big) \quad \text{and} \quad a - b = v - u.$$

By homogeneity of the norm (6.9(b)), $\|a + b\|^2 = 4\big\|w - \frac{1}{2}(u + v)\big\|^2$ and $\|a - b\|^2 = \|u - v\|^2$, so $\|a + b\|^2 + \|a - b\|^2 = 2(\|a\|^2 + \|b\|^2)$ reads

$$4\Big\|w - \tfrac{1}{2}(u + v)\Big\|^2 + \|u - v\|^2 = 2\big(\|w - u\|^2 + \|w - v\|^2\big).$$

Dividing by 4 and rearranging gives the asserted identity.

Problem 6A.32. Suppose that E is a subset of V with the property that $u, v \in E$ implies $\frac{1}{2}(u+v) \in E$. Let $w \in V$. Show that there is at most one point in E that is closest to w. In other words, show that there is at most one $u \in E$ such that

$$\|w - u\| \le \|w - x\|$$

for all $x \in E$.

Solution. Suppose $u_1, u_2 \in E$ both minimize the distance to w; the midpoint of u_1 and u_2 then beats them unless they coincide. Indeed, taking $x = u_2$ in the minimizing property of u_1 and $x = u_1$ in that of u_2 gives $\|w - u_1\| = \|w - u_2\|$; call this common value d. The hypothesis puts $\frac{1}{2}(u_1 + u_2)$ in E, so the minimizing property of u_1 applied to it gives $d^2 \le \left\|w - \frac{1}{2}(u_1 + u_2)\right\|^2$, while Exercise 31 (with $u = u_1$, $v = u_2$) evaluates that right side:

$$d^2 \le \left\|w - \tfrac{1}{2}(u_1 + u_2)\right\|^2 = \frac{d^2 + d^2}{2} - \frac{\|u_1 - u_2\|^2}{4} = d^2 - \frac{\|u_1 - u_2\|^2}{4}.$$

Hence $\|u_1 - u_2\|^2 \le 0$, so $\|u_1 - u_2\| = 0$ and $u_1 = u_2$ by 6.9(a).

Problem 6A.33. Suppose f, g are differentiable functions from $\mathbf{R}$ to $\mathbf{R}^n$.
(a) Show that

$$\langle f(t), g(t)\rangle' = \langle f'(t), g(t)\rangle + \langle f(t), g'(t)\rangle.$$

(b) Suppose c is a positive number and $\|f(t)\| = c$ for every $t \in \mathbf{R}$. Show that $\langle f'(t), f(t)\rangle = 0$ for every $t \in \mathbf{R}$.
(c) Interpret the result in (b) geometrically in terms of the tangent vector to a curve lying on a sphere in $\mathbf{R}^n$ centered at the origin.
A function $f\colon \mathbf{R} \to \mathbf{R}^n$ is called differentiable if there exist differentiable functions $f_1, \dots, f_n$ from $\mathbf{R}$ to $\mathbf{R}$ such that $f(t) = (f_1(t), \dots, f_n(t))$ for each $t \in \mathbf{R}$. Furthermore, for each $t \in \mathbf{R}$, the derivative $f'(t) \in \mathbf{R}^n$ is defined by $f'(t) = ({f_1}'(t), \dots, {f_n}'(t))$.

Solution. (a) This is the one-variable product rule applied coordinatewise. With the Euclidean inner product of 6.3(a) on $\mathbf{R}^n$ we have $\langle f(t), g(t)\rangle = \sum_{k=1}^{n} f_k(t) g_k(t)$, a finite sum of products of differentiable real-valued functions, so

$$\begin{aligned}\langle f(t), g(t)\rangle' &= \sum_{k=1}^{n} {f_k}'(t)\, g_k(t) + \sum_{k=1}^{n} f_k(t)\, {g_k}'(t)\\ &= \langle f'(t), g(t)\rangle + \langle f(t), g'(t)\rangle,\end{aligned}$$

by the definitions of $f'(t)$ and $g'(t)$.
(b) The function $t \mapsto \langle f(t), f(t)\rangle = \|f(t)\|^2 = c^2$ is constant, so its derivative is 0; by (a) with $g = f$, and the symmetry of the inner product on the real space $\mathbf{R}^n$, that derivative equals $2\langle f'(t), f(t)\rangle$. Hence $\langle f'(t), f(t)\rangle = 0$.
(c) The curve traced by f lies on the sphere of radius c centered at the origin, and (b) says its tangent vector $f'(t)$ is orthogonal (6.10) to the radius vector $f(t)$ at every point.

Problem 6A.34. Use inner products to prove Apollonius's identity: In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$

Solution. Apollonius's identity is Exercise 31 rewritten. Label the vertices u, v, w in $\mathbf{R}^2$ (with the Euclidean inner product of 6.3(a)) so that the side of length c joins u and v and w is the opposite vertex:

$$c = \|u - v\|, \qquad a = \|w - v\|, \qquad b = \|w - u\|$$

(the identity is symmetric in a and b, so this naming loses no generality). The midpoint of the side of length c is $\frac{1}{2}(u+v)$, since by 6.9(b) it lies on that segment at distance $\frac{1}{2}\|u-v\|$ from each endpoint; hence $d = \left\|w - \frac{1}{2}(u+v)\right\|$. Exercise 31 now reads

$$d^2 = \frac{b^2+a^2}{2} - \frac{c^2}{4},$$

which rearranges to $a^2+b^2 = \frac{1}{2}c^2 + 2d^2$.
Method (2): translate so the midpoint is the origin, that is, $v=-u$; then $c = \|2u\| = 2\|u\|$, $d = \|w\|$, and the parallelogram equality 6.21 gives

$$a^2+b^2 = \|w+u\|^2 + \|w-u\|^2 = 2\big(\|w\|^2+\|u\|^2\big) = 2d^2 + \tfrac{1}{2}c^2.$$

Problem 6A.35. Fix a positive integer n. The Laplacian Δp of a twice differentiable real-valued function p on $\mathbf{R}^n$ is the function on $\mathbf{R}^n$ defined by

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \cdots + \frac{\partial^2 p}{\partial x_n^2}.$$

The function p is called harmonic if $\Delta p = 0$.
A polynomial on $\mathbf{R}^n$ is a linear combination (with coefficients in $\mathbf{R}$) of functions of the form $x_1^{m_1} \cdots x_n^{m_n}$, where $m_1, \dots, m_n$ are nonnegative integers.
Suppose q is a polynomial on $\mathbf{R}^n$. Prove that there exists a harmonic polynomial p on $\mathbf{R}^n$ such that $p(x) = q(x)$ for every $x \in \mathbf{R}^n$ with $\|x\| = 1$.
The only fact about harmonic functions that you need for this exercise is that if p is a harmonic function on $\mathbf{R}^n$ and $p(x) = 0$ for all $x \in \mathbf{R}^n$ with $\|x\| = 1$, then $p = 0$.
Hint: A reasonable guess is that the desired harmonic polynomial p is of the form $q + (1-\|x\|^2)r$ for some polynomial r. Prove that there is a polynomial r on $\mathbf{R}^n$ such that $q + (1-\|x\|^2)r$ is harmonic by defining an operator T on a suitable vector space by

$$Tr = \Delta\big((1-\|x\|^2)r\big)$$

and then showing that T is injective and hence surjective.

Solution. Take $p = q + (1-\|x\|^2)r$ with r produced below. Since $1-\|x\|^2 = 1 - x_1^2 - \cdots - x_n^2$ is itself a polynomial, p is a polynomial, and $p = q$ on the unit sphere because that factor vanishes there; only $\Delta p = 0$ needs work.
For $k \geq 0$ let $\mathcal{P}_k$ be the span of the monomials $x_1^{m_1} \cdots x_n^{m_n}$ with $m_1 + \cdots + m_n \leq k$, and let $\mathcal{P}_k = \{0\}$ for $k < 0$; each $\mathcal{P}_k$ is finite-dimensional (finitely many such monomials) and every polynomial lies in some $\mathcal{P}_k$. Fix m with $q \in \mathcal{P}_m$. Since

$$\frac{\partial^2}{\partial x_j^2}\big(x_1^{m_1} \cdots x_n^{m_n}\big) = m_j(m_j-1)\, x_1^{m_1} \cdots x_j^{m_j-2} \cdots x_n^{m_n}$$

is 0 when $m_j \leq 1$ and otherwise a scalar times a monomial of total degree two lower, the linear map Δ sends $\mathcal{P}_k$ into $\mathcal{P}_{k-2}$; and multiplication by $1-\|x\|^2$ sends $\mathcal{P}_k$ into $\mathcal{P}_{k+2}$. Hence

$$Tr = \Delta\big((1-\|x\|^2)r\big)$$

defines an operator T on the finite-dimensional space $\mathcal{P}_{m-2}$, linear as a composition of those two linear maps.
T is injective. If $Tr = 0$, then $h = (1-\|x\|^2)r$ is a polynomial with $\Delta h = 0$ and $h = 0$ on the unit sphere, so $h = 0$ by the fact quoted in the statement. Thus $r(x) = 0$ whenever $\|x\| \neq 1$, and r, being a polynomial and hence continuous, vanishes on the unit sphere too (for $\|x\| = 1$, the points $(1+1/k)x$ converge to x and have norm $\neq 1$). So $\operatorname{null} T = \{0\}$, giving injectivity by 3.15 and therefore surjectivity by 3.65.

Since $q \in \mathcal{P}_m$ gives $-\Delta q \in \mathcal{P}_{m-2}$, surjectivity supplies $r \in \mathcal{P}_{m-2}$ with $Tr = -\Delta q$, and then linearity of Δ gives

$$\Delta p = \Delta q + \Delta\big((1 - \|x\|^2)r\big) = \Delta q - \Delta q = 0.$$

6.2 Exercises 6B

Problem 6B.1. Suppose $e_1, \ldots, e_m$ is a list of vectors in V such that

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \ldots, a_m \in \mathbb{F}$. Show that $e_1, \ldots, e_m$ is an orthonormal list.
This exercise provides a converse to 6.24.

Solution. Apply the hypothesis with one, then two, nonzero coefficients. Taking $a_k = 1$ and all other coefficients 0 gives $\|e_k\|^2 = 1$ for each k.
Now fix $j \neq k$, let $a \in \mathbb{F}$, and take coefficient 1 on e_j, coefficient a on e_k, and 0 elsewhere. The hypothesis gives $\|e_j + ae_k\|^2 = 1 + |a|^2$, while expanding by 6.6 and using $\|e_j\| = \|e_k\| = 1$ and $\langle e_k, e_j\rangle = \overline{\langle e_j, e_k\rangle}$ gives

$$\|e_j + ae_k\|^2 = 1 + |a|^2 + \overline{a}\langle e_j, e_k\rangle + \overline{\overline{a}\langle e_j, e_k\rangle} = 1 + |a|^2 + 2\operatorname{Re}\big(\overline{a}\langle e_j, e_k\rangle\big).$$

Hence $\operatorname{Re}\big(\overline{a}\langle e_j, e_k\rangle\big) = 0$ for every $a \in \mathbb{F}$.
(i) $\mathbb{F} = \mathbb{R}$: taking $a = 1$ gives $\langle e_j, e_k\rangle = \operatorname{Re}\langle e_j, e_k\rangle = 0$.
(ii) $\mathbb{F} = \mathbb{C}$: write $\langle e_j, e_k\rangle = x + iy$ with $x, y \in \mathbb{R}$. Then $a = 1$ gives $x = 0$, and $a = i$ gives

$$0 = \operatorname{Re}\big(-i(x + iy)\big) = \operatorname{Re}(y - ix) = y.$$

So the list is orthonormal.

Problem 6B.2. (a) Suppose $\theta \in \mathbb{R}$. Show that both

$$(\cos\theta, \sin\theta),\ (-\sin\theta, \cos\theta) \qquad \text{and} \qquad (\cos\theta, \sin\theta),\ (\sin\theta, -\cos\theta)$$

are orthonormal bases of $\mathbb{R}^2$.
(b) Show that each orthonormal basis of $\mathbb{R}^2$ is of the form given by one of the two possibilities in (a).

Solution. Throughout, $\mathbb{R}^2$ carries the Euclidean inner product.
(a) By $\cos^2\theta + \sin^2\theta = 1$ each of the three vectors has norm 1, and

$$\langle(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\rangle = 0 = \langle(\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)\rangle.$$

(Check!) So both lists are orthonormal of length $2 = \dim \mathbb{R}^2$, hence orthonormal bases by 6.28.
(b) Let $u = (a, b)$, $w = (c, d)$ be an orthonormal basis of $\mathbb{R}^2$ (every basis has length 2), so that

$$a^2 + b^2 = 1, \qquad c^2 + d^2 = 1, \qquad ac + bd = 0.$$

Since $|a| \leq 1$, set $\theta = \arccos a$ if $b \geq 0$ and $\theta = -\arccos a$ if $b < 0$; then $\cos\theta = a$ and $\sin\theta = \pm\sqrt{1 - a^2} = b$, so $u = (\cos\theta, \sin\theta)$. Put $s = bc - ad$. Multiplying $ac + bd = 0$ by a gives $abd = -a^2c$, and multiplying it by b gives $abc = -b^2d$, so

$$\begin{aligned} bs &= b^2c - abd = (a^2 + b^2)c = c, \\ as &= abc - a^2d = -(a^2 + b^2)d = -d. \end{aligned}$$

Thus $w = s(b, -a)$, and $1 = c^2 + d^2 = s^2(a^2 + b^2) = s^2$, so $s = \pm 1$.
(i) $s = -1$: then $w = (-b, a) = (-\sin\theta, \cos\theta)$, the first list in (a).
(ii) $s = 1$: then $w = (b, -a) = (\sin\theta, -\cos\theta)$, the second list in (a).

Problem 6B.3. Suppose $e_1, \ldots, e_m$ is an orthonormal list in V and $v \in V$. Prove that

$$\|v\|^2 = \big|\langle v, e_1\rangle\big|^2 + \cdots + \big|\langle v, e_m\rangle\big|^2 \iff v \in \operatorname{span}(e_1, \ldots, e_m).$$

Solution. Both sides of the equivalence say that $w = 0$, where, as in the proof of Bessel's inequality 6.26,

$$u = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m, \qquad w = v - u.$$

Orthonormality gives $\langle w, e_k\rangle = \langle v, e_k\rangle - \langle v, e_k\rangle = 0$ for each k, so by linearity in the first slot w is orthogonal to every vector of $\operatorname{span}(e_1, \ldots, e_m)$, in particular to u. Hence the Pythagorean theorem 6.12 and 6.24 give

$$\|v\|^2 = \|u\|^2 + \|w\|^2 = \big|\langle v, e_1\rangle\big|^2 + \cdots + \big|\langle v, e_m\rangle\big|^2 + \|w\|^2,$$

so the asserted equality holds if and only if $w = 0$, that is, if and only if $v = u$.
If the equality holds, then $v = u \in \operatorname{span}(e_1, \ldots, e_m)$. Conversely, if $v \in \operatorname{span}(e_1, \ldots, e_m)$, then $w = v - u$ lies in that span too, and w is orthogonal to it, so $\langle w, w\rangle = 0$ and $w = 0$.

Problem 6B.4. Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \ \frac{\cos x}{\sqrt{\pi}}, \ \frac{\cos 2x}{\sqrt{\pi}}, \ \ldots, \ \frac{\cos nx}{\sqrt{\pi}}, \ \frac{\sin x}{\sqrt{\pi}}, \ \frac{\sin 2x}{\sqrt{\pi}}, \ \ldots, \ \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} fg.$$

Hint: The following formulas should help.

$$(\sin x)(\cos y) = \frac{\sin(x-y) + \sin(x+y)}{2}$$
$$(\sin x)(\sin y) = \frac{\cos(x-y) - \cos(x+y)}{2}$$
$$(\cos x)(\cos y) = \frac{\cos(x-y) + \cos(x+y)}{2}$$

Solution. Everything reduces to the two vanishing integrals

$$\int_{-\pi}^{\pi} \sin(mx)\,dx = 0, \qquad \int_{-\pi}^{\pi} \cos(mx)\,dx = 0 \ \text{ if } m \neq 0,$$

valid for every integer m; call them $(*)$. Indeed, for $m = 0$ the first integrand vanishes identically, while for $m \neq 0$,

$$\left[\frac{-\cos(mx)}{m}\right]_{-\pi}^{\pi} = 0 \qquad \text{and} \qquad \left[\frac{\sin(mx)}{m}\right]_{-\pi}^{\pi} = 0$$

because cosine is even and $\sin(m\pi) = 0$. Below $j, k \in \{1, \ldots, n\}$, so $j + k \neq 0$, and $j - k = 0$ exactly when $j = k$.
Norms. First $\|1/\sqrt{2\pi}\|^2 = \int_{-\pi}^{\pi} (2\pi)^{-1} dx = 1$. Next, the third and second hint formulas with both x and y taken to be kx give $\cos^2(kx) = \frac{1}{2}(1 + \cos 2kx)$ and $\sin^2(kx) = \frac{1}{2}(1 - \cos 2kx)$, so by $(*)$ with $m = 2k \neq 0$,

$$\left\|\frac{\cos kx}{\sqrt{\pi}}\right\|^2 = \frac{1}{\pi} \cdot \frac{2\pi}{2} = 1 = \left\|\frac{\sin kx}{\sqrt{\pi}}\right\|^2.$$

Orthogonality. The inner products of $1/\sqrt{2\pi}$ with $\cos(kx)/\sqrt{\pi}$ and with $\sin(kx)/\sqrt{\pi}$ are multiples of $\int_{-\pi}^{\pi} \cos(kx)\,dx$ and $\int_{-\pi}^{\pi} \sin(kx)\,dx$, hence 0 by $(*)$. For the remaining pairs the hint formulas turn each inner product into a combination of two integrals covered by $(*)$. By the first hint formula, for all j, k (including $j = k$), since $(*)$ kills every sine integral,

$$\left\langle \frac{\cos jx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} \sin\big((k-j)x\big)dx + \int_{-\pi}^{\pi} \sin\big((k+j)x\big)dx\right) = 0,$$

and by the third and second hint formulas, for $j \neq k$ (so that both $j-k$ and $j+k$ are nonzero),

$$\left\langle \frac{\cos jx}{\sqrt{\pi}}, \frac{\cos kx}{\sqrt{\pi}} \right\rangle = \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} \cos\big((j-k)x\big)dx + \int_{-\pi}^{\pi} \cos\big((j+k)x\big)dx\right) = 0,$$
$$\left\langle \frac{\sin jx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} \cos\big((j-k)x\big)dx - \int_{-\pi}^{\pi} \cos\big((j+k)x\big)dx\right) = 0.$$

So all $2n+1$ functions have norm 1 and are pairwise orthogonal.

Problem 6B.5. Suppose $f\colon [-\pi, \pi] \to \mathbb{R}$ is continuous. For each nonnegative integer k, define

$$a_k = \frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\cos(kx)\,dx \qquad \text{and} \qquad b_k = \frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f(x)\sin(kx)\,dx.$$

Prove that

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty}\big(a_k^2 + b_k^2\big) \le \int_{-\pi}^{\pi} f^2.$$

The inequality above is actually an equality for all continuous functions $f\colon [-\pi, \pi] \to \mathbb{R}$. However, proving that this inequality is an equality involves Fourier series techniques beyond the scope of this book.

Solution. This is Bessel's inequality 6.26 for the orthonormal list of Exercise 6B.4, followed by $n \to \infty$. Work in $V = C[-\pi, \pi]$ with $\langle g, h\rangle = \int_{-\pi}^{\pi} gh$, an inner product by the verification of 6.3(c) (stated there for $[-1, 1]$; the one nonformal point is that a continuous $g \neq 0$ is bounded away from 0 on some subinterval, so $\int_{-\pi}^{\pi} g^2 > 0$). Thus $\|f\|^2 = \int_{-\pi}^{\pi} f^2$.
Fix a positive integer n. By Exercise 6B.4 the list $e_0 = 1/\sqrt{2\pi}$, $c_k = \cos(kx)/\sqrt{\pi}$, $s_k = \sin(kx)/\sqrt{\pi}$ for $k = 1, \dots, n$ is orthonormal in V, and

$$\langle f, e_0\rangle = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{\pi}}\int_{-\pi}^{\pi} f = \frac{a_0}{\sqrt{2}}, \qquad \langle f, c_k\rangle = a_k, \qquad \langle f, s_k\rangle = b_k,$$

since $\cos(0 \cdot x) = 1$. So 6.26 applied to f and this list gives

$$S_n = \frac{a_0^2}{2} + \sum_{k=1}^{n}\big(a_k^2 + b_k^2\big) \le \|f\|^2 = \int_{-\pi}^{\pi} f^2.$$

The terms $a_k^2 + b_k^2$ are nonnegative, so (S_n) is increasing and bounded above by $\int_{-\pi}^{\pi} f^2$; hence it converges, and the sum of the infinite series is $\lim_{n\to\infty} S_n = \sup_n S_n \le \int_{-\pi}^{\pi} f^2$.

Problem 6B.6. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V.
(a) Prove that if $v_1, \dots, v_n$ are vectors in V such that

$$\|e_k - v_k\| < \frac{1}{\sqrt{n}}$$

for each k, then $v_1, \dots, v_n$ is a basis of V.

(b) Show that there exist $v_1, \ldots, v_n \in V$ such that

$$\|e_k - v_k\| \le \frac{1}{\sqrt{n}}$$

for each k, but $v_1, \ldots, v_n$ is not linearly independent.
This exercise states in (a) that an appropriately small perturbation of an orthonormal basis is a basis. Then (b) shows that the number $1/\sqrt{n}$ on the right side of the inequality in (a) cannot be higher.

Solution. (a) It is enough to prove $v_1, \ldots, v_n$ linearly independent, since $\dim V = n$ and a linearly independent list of that length is a basis by 2.38. Suppose $a_1 v_1 + \cdots + a_n v_n = 0$ with not all a_k equal to 0, and set $c = \left(\sum_{k=1}^n |a_k|^2\right)^{1/2} > 0$. Subtracting that relation,

$$a_1 e_1 + \cdots + a_n e_n = \sum_{k=1}^n a_k (e_k - v_k).$$

Now chain 6.24 (for the left side), the triangle inequality 6.17 with homogeneity, the bound $|a_k|\,\|e_k - v_k\| \le |a_k|/\sqrt{n}$ (strict for the indices with $a_k \ne 0$, of which there is at least one), and the Cauchy–Schwarz inequality 6.14 in $\mathbb{R}^n$ applied to $(|a_1|, \ldots, |a_n|)$ and $(1, \ldots, 1)$:

$$\begin{aligned} c = \|a_1 e_1 + \cdots + a_n e_n\| &\le \sum_{k=1}^n |a_k|\,\|e_k - v_k\| \\ &< \frac{1}{\sqrt{n}} \sum_{k=1}^n |a_k| \le \frac{1}{\sqrt{n}} \cdot \sqrt{n}\, c = c, \end{aligned}$$

which is impossible. Hence every a_k equals 0.
(b) Take $e = e_1 + \cdots + e_n$ and $v_k = e_k - \frac{1}{n} e$. By 6.24 (all coefficients equal 1) we have $\|e\| = \sqrt{n}$, so

$$\|e_k - v_k\| = \left\|\tfrac{1}{n} e\right\| = \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}} \qquad \text{for each } k,$$

while $v_1 + \cdots + v_n = e - n \cdot \frac{1}{n} e = 0$ is a linear combination equal to 0 with all coefficients 1. So $v_1, \ldots, v_n$ is not linearly independent.

Problem 6B.7. Suppose $T \in \mathcal{L}(\mathbb{R}^3)$ has an upper-triangular matrix with respect to the basis $(1,0,0)$, $(1,1,1)$, $(1,1,2)$. Find an orthonormal basis of $\mathbb{R}^3$ with respect to which T has an upper-triangular matrix.

Solution. The answer is

$$e_1 = (1,0,0), \qquad e_2 = \left(0, \tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right), \qquad e_3 = \left(0, -\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right),$$

the output of the Gram–Schmidt procedure 6.32 applied to $v_1 = (1,0,0)$, $v_2 = (1,1,1)$, $v_3 = (1,1,2)$ with the Euclidean inner product: using $\langle v_2, f_1\rangle = 1$, $\langle v_3, f_1\rangle = 1$, $\langle v_3, f_2\rangle = 3$ and $e_k = f_k/\|f_k\|$,

$$\begin{aligned} f_1 &= (1,0,0), & \|f_1\|^2 &= 1, \\ f_2 &= (1,1,1) - (1,0,0) = (0,1,1), & \|f_2\|^2 &= 2, \\ f_3 &= (1,1,2) - (1,0,0) - \tfrac{3}{2}(0,1,1) = \left(0, -\tfrac{1}{2}, \tfrac{1}{2}\right), & \|f_3\|^2 &= \tfrac{1}{2}. \end{aligned}$$

This is an orthonormal list of length $3 = \dim \mathbb{R}^3$, hence an orthonormal basis by 6.28.
It has the required property because 6.32 gives $\operatorname{span}(e_1, \ldots, e_k) = \operatorname{span}(v_1, \ldots, v_k)$ for $k = 1, 2, 3$: these subspaces are invariant under T by 5.39 (since T has an upper-triangular matrix with respect to v_1, v_2, v_3), so 5.39 applied to the basis e_1, e_2, e_3 makes the matrix of T with respect to it upper triangular.

Problem 6B.8. Make $\mathcal{P}_2(\mathbf{R})$ into an inner product space by defining $\langle p,q\rangle = \int_0^1 pq$ for all $p,q \in \mathcal{P}_2(\mathbf{R})$.
(a) Apply the Gram–Schmidt procedure to the basis $1, x, x^2$ to produce an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.
(b) The differentiation operator (the operator that takes p to p') on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to the basis $1, x, x^2$, which is not an orthonormal basis. Find the matrix of the differentiation operator on $\mathcal{P}_2(\mathbf{R})$ with respect to the orthonormal basis produced in (a) and verify that this matrix is upper triangular, as expected from the proof of 6.37.

Solution. (a) The Gram–Schmidt procedure 6.32 applied to $v_1 = 1$, $v_2 = x$, $v_3 = x^2$ produces

$$e_1 = 1, \qquad e_2 = \sqrt{3}\,(2x-1), \qquad e_3 = \sqrt{5}\,(6x^2-6x+1).$$

Indeed $f_1 = 1$ with $\|f_1\|^2 = \int_0^1 1\,dt = 1$; since $\langle v_2, f_1\rangle = \int_0^1 t\,dt = \frac{1}{2}$,

$$f_2 = x - \tfrac{1}{2}, \qquad \|f_2\|^2 = \int_0^1 \left(t-\tfrac{1}{2}\right)^2 dt = \tfrac{1}{12},$$

so $\|f_2\| = 1/(2\sqrt{3})$ and $e_2 = 2\sqrt{3}\,(x-\frac{1}{2})$. Since $\langle v_3, f_1\rangle = \frac{1}{3}$ and $\langle v_3, f_2\rangle = \frac{1}{4} - \frac{1}{6} = \frac{1}{12}$,

$$f_3 = x^2 - \tfrac{1}{3} - \left(x - \tfrac{1}{2}\right) = x^2 - x + \tfrac{1}{6},$$

and expanding $\left(t^2 - t + \frac{1}{6}\right)^2 = t^4 - 2t^3 + \frac{4}{3}t^2 - \frac{1}{3}t + \frac{1}{36}$ gives

$$\|f_3\|^2 = \tfrac{1}{5} - \tfrac{1}{2} + \tfrac{4}{9} - \tfrac{1}{6} + \tfrac{1}{36} = \frac{1}{180},$$

so $\|f_3\| = 1/(6\sqrt{5})$ and $e_3 = 6\sqrt{5}f_3$. This orthonormal list has length $3 = \dim \mathcal{P}_2(\mathbf{R})$, hence is an orthonormal basis by 6.28.
(b) The matrix is

$$\mathcal{M}\big(D,(e_1,e_2,e_3)\big) = \begin{pmatrix} 0 & 2\sqrt{3} & 0 \\ 0 & 0 & 2\sqrt{15} \\ 0 & 0 & 0 \end{pmatrix},$$

which is upper triangular. Indeed $De_1 = 0$, $De_2 = 2\sqrt{3} = 2\sqrt{3}\,e_1$, and

$$De_3 = 6\sqrt{5}\,(2x-1) = \frac{6\sqrt{5}}{\sqrt{3}}\,e_2 = 2\sqrt{15}\,e_2.$$

This is what the proof of 6.37 predicts: $\operatorname{span}(v_1,\dots,v_k)$ is D-invariant by 5.39, it equals $\operatorname{span}(e_1,\dots,e_k)$ by 6.32, and 5.39 then applies to e_1, e_2, e_3.

Problem 6B.9. Suppose $e_1,\dots,e_m$ is the result of applying the Gram–Schmidt procedure to a linearly independent list $v_1,\dots,v_m$ in V. Prove that $\langle v_k, e_k\rangle > 0$ for each $k = 1,\dots,m$.

Solution. In fact $\langle v_k, e_k\rangle = \|f_k\| > 0$, where $f_1,\dots,f_m$ are the vectors of the Gram–Schmidt formulas 6.32 and $e_k = f_k/\|f_k\|$; here $f_k \neq 0$ for each k, as shown in the proof of 6.32 from the linear independence of $v_1,\dots,v_m$. Rearranging the formula for f_k (and reading it as $v_1 = f_1$ when $k=1$) gives $v_k = f_k + u_k$ with

$$u_k = \sum_{j=1}^{k-1} \frac{\langle v_k, f_j\rangle}{\|f_j\|^2} f_j \in \operatorname{span}(f_1,\dots,f_{k-1}) = \operatorname{span}(e_1,\dots,e_{k-1}),$$

the last equality because each e_j is a nonzero scalar multiple of f_j. Since $e_1,\dots,e_m$ is orthonormal by

6.32, the vector e_k is orthogonal to that span, so $\langle u_k, e_k\rangle = 0$ and

$$\langle v_k, e_k\rangle = \langle f_k, e_k\rangle = \frac{1}{\|f_k\|}\langle f_k, f_k\rangle = \|f_k\| > 0,$$

the middle step by homogeneity in the second slot (the conjugate is harmless, the scalar $1/\|f_k\|$ being real).

Problem 6B.10. Suppose $v_1, \ldots, v_m$ is a linearly independent list in V. Explain why the orthonormal list produced by the formulas of the Gram–Schmidt procedure (6.32) is the only orthonormal list $e_1, \ldots, e_m$ in V such that $\langle v_k, e_k\rangle > 0$ and $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$ for each $k = 1, \ldots, m$. [The result in this exercise is used in the proof of 7.58.]

Solution. The Gram–Schmidt list $e_1, \ldots, e_m$ of 6.32 has both properties (the span identities by 6.32, and $\langle v_k, e_k\rangle > 0$ by Exercise 6B.9), and any orthonormal list $g_1, \ldots, g_m$ with both properties equals it, by induction on k.
Let $1 \le k \le m$ and suppose $g_j = e_j$ for all $j < k$. Since

$$\operatorname{span}(g_1, \ldots, g_k) = \operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k),$$

the vector g_k lies in $\operatorname{span}(e_1, \ldots, e_k)$, of which $e_1, \ldots, e_k$ is an orthonormal basis, so 6.30(a) applied inside that subspace gives

$$g_k = \langle g_k, e_1\rangle e_1 + \cdots + \langle g_k, e_k\rangle e_k.$$

For $j < k$ the induction hypothesis gives $e_j = g_j$, and $\langle g_k, g_j\rangle = 0$ by orthonormality of $g_1, \ldots, g_m$; so the expansion collapses to $g_k = a e_k$ with $a = \langle g_k, e_k\rangle$, and $1 = \|g_k\| = |a|$. Finally

$$\langle v_k, g_k\rangle = \overline{a}\,\langle v_k, e_k\rangle,$$

where $\langle v_k, g_k\rangle > 0$ by hypothesis and $\langle v_k, e_k\rangle > 0$ by 6B.9, so $\overline{a} > 0$; with $|a| = 1$ this forces $a = 1$ and $g_k = e_k$. (For $k = 1$ the induction hypothesis is vacuous and the same argument applies.)

Problem 6B.11. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that $p\big(\frac{1}{2}\big) = \int_0^1 pq$ for every $p \in \mathcal{P}_2(\mathbf{R})$.

Solution. The answer is $q(x) = -15x^2 + 15x - \frac{3}{2}$.
Give $\mathcal{P}_2(\mathbf{R})$ the inner product $\langle p, q\rangle = \int_0^1 pq$. The functional $\varphi(p) = p\big(\frac{1}{2}\big)$ is linear on this finite-dimensional space, so the Riesz representation theorem 6.42 supplies the required q, and 6.43 computes it as $q = \varphi(e_1)e_1 + \varphi(e_2)e_2 + \varphi(e_3)e_3$ for any orthonormal basis (no conjugates, the field being $\mathbf{R}$). Take the basis of Exercise 6B.8(a), namely $e_1 = 1$, $e_2 = \sqrt{3}\,(2x - 1)$, $e_3 = \sqrt{5}\,(6x^2 - 6x + 1)$, and evaluate at $x = \frac{1}{2}$:

$$\varphi(e_1) = 1, \qquad \varphi(e_2) = \sqrt{3}\,(1 - 1) = 0, \qquad \varphi(e_3) = \sqrt{5}\,\big(\tfrac{3}{2} - 3 + 1\big) = -\frac{\sqrt{5}}{2}.$$

Hence

$$q = 1 - \frac{\sqrt{5}}{2} \cdot \sqrt{5}\,(6x^2 - 6x + 1) = -15x^2 + 15x - \tfrac{3}{2}.$$

Problem 6B.12. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x)\cos(\pi x)\,dx = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

Solution. The answer is $q(x) = \dfrac{12(1-2x)}{\pi^2}$.
Give $\mathcal{P}_2(\mathbf{R})$ the inner product $\langle p, q\rangle = \int_0^1 pq$. The functional $\varphi(p) = \int_0^1 p(x)\cos(\pi x)\,dx$ is linear on this finite-dimensional space, so 6.42 supplies the required q and 6.43 gives $q = \varphi(e_1)e_1 + \varphi(e_2)e_2 + \varphi(e_3)e_3$ (no conjugates, the field being $\mathbf{R}$) for the orthonormal basis $e_1 = 1$, $e_2 = \sqrt{3}\,(2x-1)$, $e_3 = \sqrt{5}\,(6x^2 - 6x+1)$ of Exercise 6B.8(a). The integrals needed are

$$\int_0^1 \cos(\pi x)\,dx = \left[\frac{\sin(\pi x)}{\pi}\right]_0^1 = 0, \qquad \int_0^1 \sin(\pi x)\,dx = \left[\frac{-\cos(\pi x)}{\pi}\right]_0^1 = \frac{2}{\pi},$$
$$\int_0^1 x\sin(\pi x)\,dx = \frac{1}{\pi} + \frac{1}{\pi}\cdot 0 = \frac{1}{\pi}, \qquad \int_0^1 x\cos(\pi x)\,dx = 0 - \frac{1}{\pi}\cdot\frac{2}{\pi} = -\frac{2}{\pi^2},$$
$$\int_0^1 x^2\cos(\pi x)\,dx = 0 - \frac{2}{\pi}\cdot\frac{1}{\pi} = -\frac{2}{\pi^2},$$

the last three by parts. Hence

$$\varphi(e_1) = 0, \qquad \varphi(e_2) = \sqrt{3}\left(-\frac{4}{\pi^2} - 0\right) = -\frac{4\sqrt{3}}{\pi^2}, \qquad \varphi(e_3) = 0,$$

the last because $6\left(-\frac{2}{\pi^2}\right) - 6\left(-\frac{2}{\pi^2}\right) + 0 = 0$, and therefore

$$q = -\frac{4\sqrt{3}}{\pi^2}\cdot\sqrt{3}\,(2x-1) = \frac{12}{\pi^2}\,(1-2x).$$

Problem 6B.13. Show that a list $v_1, \ldots, v_m$ of vectors in V is linearly dependent if and only if the Gram–Schmidt formula in 6.32 produces $f_k = 0$ for some $k \in \{1, \ldots, m\}$.
[This exercise gives an alternative to Gaussian elimination techniques for determining whether a list of vectors in an inner product space is linearly dependent.]

Solution. The formula 6.32 for f_k makes sense only while $f_1, \ldots, f_{k-1}$ are all nonzero, since $\|f_j\|^2$ sits in a denominator; so the procedure produces $f_1 = v_1, f_2, \ldots$ in turn and halts at the first index with output 0, and the assertion "produces $f_k = 0$ for some $k \le m$" means that it halts at or before index m.
Two facts hold whenever $f_1, \ldots, f_k$ are defined in this sense, both by induction on k and neither using linear independence of $v_1, \ldots, v_m$.
(i) The list $f_1, \ldots, f_k$ is orthogonal: for $j < k$, the inductive hypothesis kills every term of the sum but the $i = j$ one, so

$$\langle f_k, f_j\rangle = \langle v_k, f_j\rangle - \sum_{i=1}^{k-1}\frac{\langle v_k, f_i\rangle}{\|f_i\|^2}\langle f_i, f_j\rangle = \langle v_k, f_j\rangle - \langle v_k, f_j\rangle = 0.$$

(ii) $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(f_1, \ldots, f_k)$: here $f_1 = v_1$, while the formula puts f_k in $\operatorname{span}(v_k, f_1, \ldots, f_{k-1})$ and, rearranged, puts v_k in $\operatorname{span}(f_1, \ldots, f_k)$.
Suppose first that $f_k = 0$, with k least. If $k = 1$ then $v_1 = 0$ and the list is linearly dependent. If $k > 1$, the formula rearranges to

$$v_k = \sum_{j=1}^{k-1}\frac{\langle v_k, f_j\rangle}{\|f_j\|^2}f_j \in \operatorname{span}(f_1, \ldots, f_{k-1}) = \operatorname{span}(v_1, \ldots, v_{k-1})$$

by (ii), so $v_1, \ldots, v_m$ is linearly dependent (move v_k across, with coefficient -1).
Conversely, suppose $f_1, \ldots, f_m$ are all defined and nonzero. By (i) they form an orthogonal list of nonzero vectors, so the $e_k = f_k/\|f_k\|$ form an orthonormal, hence by 6.25 linearly independent, list; thus $\operatorname{span}(f_1, \ldots, f_m)$ has dimension m, and by (ii) so does $U = \operatorname{span}(v_1, \ldots, v_m)$. A list of length $\dim U$ spanning U is a basis of U by 2.42, so $v_1, \ldots, v_m$ is linearly independent.

Problem 6B.14. Suppose V is a real inner product space and $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $e_1, \ldots, e_m$ of vectors in V such that

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$$

for all $k \in \{1, \ldots, m\}$.

Solution. The lists in question are exactly $\varepsilon_1 g_1, \ldots, \varepsilon_m g_m$ with $\varepsilon_1, \ldots, \varepsilon_m \in \{1, -1\}$, where $g_1, \ldots, g_m$ is the Gram–Schmidt list 6.32 for $v_1, \ldots, v_m$; there are 2^m of these.
Each such list works: $\|\varepsilon_k g_k\| = 1$ and $\langle \varepsilon_j g_j, \varepsilon_k g_k\rangle = \varepsilon_j \varepsilon_k \langle g_j, g_k\rangle = 0$ for $j \neq k$, so the list is orthonormal, and since each $\varepsilon_k \neq 0$,

$$\operatorname{span}(\varepsilon_1 g_1, \ldots, \varepsilon_k g_k) = \operatorname{span}(g_1, \ldots, g_k) = \operatorname{span}(v_1, \ldots, v_k),$$

the last equality by 6.32.
There are no others. Let $e_1, \ldots, e_m$ be orthonormal with the stated span property and fix k; then

$$\operatorname{span}(e_1, \ldots, e_k) = \operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(g_1, \ldots, g_k),$$

so e_k lies in the last span, of which $g_1, \ldots, g_k$ is an orthonormal basis, and 6.30(a) applied inside it expands e_k in $g_1, \ldots, g_k$. For $j < k$ we have $g_j \in \operatorname{span}(g_1, \ldots, g_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1})$, to which e_k is orthogonal, so $\langle e_k, g_j\rangle = 0$ (vacuous when $k = 1$) and the expansion collapses to $e_k = a_k g_k$ with $a_k = \langle e_k, g_k\rangle \in \mathbf{R}$. Then $1 = \|e_k\| = |a_k|$, and the field being $\mathbf{R}$ forces $a_k = \pm 1$.
Finally the 2^m sign choices give distinct lists: if two sign vectors differ at index k, the k-th entries differ by $\pm 2 g_k \neq 0$. So these lists are in bijection with $\{1, -1\}^m$.

Problem 6B.15. Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle u, v\rangle_1 = 0$ if and only if $\langle u, v\rangle_2 = 0$. Prove that there is a positive number c such that $\langle u, v\rangle_1 = c\langle u, v\rangle_2$ for every $u, v \in V$.
This exercise shows that if two inner products have the same pairs of orthogonal vectors, then each of the inner products is a scalar multiple of the other inner product.

Solution. Take $c = f(x)$ for any nonzero $x \in V$, where

$$f(x) = \frac{\langle x, x\rangle_1}{\langle x, x\rangle_2},$$

a positive real number by positive definiteness of both inner products; the work is to show that f is constant. (If $V = \{0\}$, both inner products vanish identically and $c = 1$ works.)
Two-dimensional subspaces. Let U be a two-dimensional subspace of V. Restricting $\langle \cdot, \cdot \rangle_2$ to U makes U an inner product space, so 6.35 gives a basis e_1, e_2 of U orthonormal for $\langle \cdot, \cdot \rangle_2$. From $\langle e_1, e_2\rangle_2 = 0$ the hypothesis gives $\langle e_1, e_2\rangle_1 = 0$, hence $\langle e_2, e_1\rangle_1 = 0$ by conjugate symmetry; and from

$$\langle e_1 + e_2, e_1 - e_2\rangle_2 = 1 - 0 + 0 - 1 = 0$$

the hypothesis gives $\langle e_1 + e_2, e_1 - e_2\rangle_1 = 0$, which expands (the two cross terms vanishing) to $\langle e_1, e_1\rangle_1 = \langle e_2, e_2\rangle_1$. Writing $c_U > 0$ for this common value and expanding $x = a_1 e_1 + a_2 e_2$, $y = b_1 e_1 + b_2 e_2$ gives

$$\langle x, y\rangle_1 = c_U(a_1 \overline{b_1} + a_2 \overline{b_2}) = c_U \langle x, y\rangle_2$$

for all $x, y \in U$; call this (i). In particular, with $y = x \neq 0$, we get $f = c_U$ on $U \setminus \{0\}$; call this (ii).
f is constant on $V \setminus \{0\}$. Let x, y be nonzero. If $y = \lambda x$ with $\lambda \neq 0$ (or $x = \lambda y$, the same argument), the factors $|\lambda|^2$ cancel in $f(y)$ and $f(y) = f(x)$; otherwise x, y is linearly independent and (ii) applied to $U = \operatorname{span}(x, y)$ gives $f(x) = f(y)$. Call the common value $c > 0$.
Conclusion. If $u = 0$ or $v = 0$, both sides vanish. If $v = \lambda u$ with $\lambda \neq 0$ (the case $u = \lambda v$ being symmetric), then $\langle u, u\rangle_1 = c\langle u, u\rangle_2$ gives

$$\langle u, v\rangle_1 = \overline{\lambda}\langle u, u\rangle_1 = c\langle u, \lambda u\rangle_2 = c\langle u, v\rangle_2.$$

If u, v is linearly independent, (i) for $U = \operatorname{span}(u, v)$ gives $\langle u, v\rangle_1 = c_U \langle u, v\rangle_2$, and $c_U = f(u) = c$ by (ii).

Problem 6B.16. Suppose V is finite-dimensional. Suppose $\langle\cdot,\cdot\rangle_1$, $\langle\cdot,\cdot\rangle_2$ are inner products on V with corresponding norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Prove that there exists a positive number c such that $\|v\|_1 \le c\|v\|_2$ for every $v \in V$.

Solution. Take $c = \sqrt{n}\, M$, where $n = \dim V$, the basis $e_1, \dots, e_n$ of V is orthonormal for $\langle\cdot,\cdot\rangle_2$ (such a basis exists by 6.35, V being finite-dimensional), and $M = \max\{\|e_1\|_1, \dots, \|e_n\|_1\} > 0$. (If $V = \{0\}$ then $c = 1$ works.)
Write $v = a_1 e_1 + \cdots + a_n e_n$. By 6.24 applied in $\langle\cdot,\cdot\rangle_2$, by the triangle inequality 6.17 and homogeneity 6.9 for $\|\cdot\|_1$, and by the Cauchy–Schwarz inequality 6.14 in $\mathbf{R}^n$ applied to $(|a_1|, \dots, |a_n|)$ and $(1, \dots, 1)$,

$$\begin{aligned} \|v\|_2^2 &= |a_1|^2 + \cdots + |a_n|^2, \\ \|v\|_1 &\le |a_1|\, \|e_1\|_1 + \cdots + |a_n|\, \|e_n\|_1 \le M\left(|a_1| + \cdots + |a_n|\right), \\ |a_1| + \cdots + |a_n| &\le \sqrt{n}\left(|a_1|^2 + \cdots + |a_n|^2\right)^{1/2} = \sqrt{n}\, \|v\|_2. \end{aligned}$$

Combining the last two lines gives $\|v\|_1 \le M\sqrt{n}\, \|v\|_2 = c\|v\|_2$.

Problem 6B.17. Suppose $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional. Prove that if T is an operator on V such that 1 is the only eigenvalue of T and $\|Tv\| \le \|v\|$ for all $v \in V$, then T is the identity operator.

Solution. Triangularize and read off the columns. The case $V = \{0\}$ is trivial, so let $n = \dim V \ge 1$; since $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional, Schur's theorem (6.38) gives an orthonormal basis $e_1, \dots, e_n$ of V for which $A = \mathcal{M}(T, (e_1, \dots, e_n))$ is upper triangular, so $Te_k = \sum_{j \le k} A_{j,k} e_j$. By 5.41 the diagonal entries are the eigenvalues of T, whence $A_{k,k} = 1$ for every k.
Fix k. By 6.24 (the e_j are orthonormal) and the hypothesis applied to $v = e_k$,

$$\begin{aligned} 1 + \sum_{j<k} |A_{j,k}|^2 = \sum_{j \le k} |A_{j,k}|^2 &= \|Te_k\|^2 \\ &\le \|e_k\|^2 = 1. \end{aligned}$$

So each $|A_{j,k}|^2 = 0$ for $j < k$, giving $Te_k = A_{k,k} e_k = e_k$. Two linear maps agreeing on a basis are equal (3.4), so $T = I$.

Problem 6B.18. Suppose $u_1, \dots, u_m$ is a linearly independent list in V. Show that there exists $v \in V$ such that $\langle u_k, v\rangle = 1$ for all $k \in \{1, \dots, m\}$.

Solution. Take v to be the Riesz vector of the functional on $U = \operatorname{span}(u_1, \dots, u_m)$ sending each u_k to 1.
Indeed, $u_1, \dots, u_m$ is a basis of U (it spans U, and it is linearly independent by hypothesis), so U is a finite-dimensional inner product space under the inner product inherited from V, and the linear map lemma (3.4) gives a linear functional $\varphi\colon U \to \mathbf{F}$ with $\varphi(u_k) = 1$ for each k. By the Riesz representation theorem (6.42), applicable since $\dim U = m < \infty$, there is $v \in U \subseteq V$ with $\varphi(u) = \langle u, v\rangle$ for all $u \in U$. Taking $u = u_k$,

$$\langle u_k, v\rangle = \varphi(u_k) = 1 \quad (k = 1, \dots, m).$$

Problem 6B.19. Suppose $v_1, \dots, v_n$ is a basis of V. Prove that there exists a basis $u_1, \dots, u_n$ of V

such that

$$\langle v_j, u_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

Solution. Take u_k to be the Riesz vector of the k-th dual-basis functional. Here $\dim V = n$, and the linear map lemma (3.4) gives, for each k, a linear functional φ_k on V with $\varphi_k(v_j) = 0$ for $j \neq k$ and $\varphi_k(v_k) = 1$; the Riesz representation theorem (6.42), applicable since V is finite-dimensional, then supplies $u_k \in V$ with $\varphi_k(u) = \langle u, u_k \rangle$ for all $u \in V$. Taking $u = v_j$ gives

$$\langle v_j, u_k \rangle = \varphi_k(v_j) = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

For linear independence, suppose $a_1 u_1 + \cdots + a_n u_n = 0$ and fix j. Pairing with v_j and using conjugate linearity in the second slot [6.6(d), 6.6(e)],

$$\begin{aligned} 0 &= \langle v_j, a_1 u_1 + \cdots + a_n u_n \rangle \\ &= \sum_{k=1}^{n} \overline{a_k} \langle v_j, u_k \rangle = \overline{a_j}. \end{aligned}$$

Thus $a_1 = \cdots = a_n = 0$, and a linearly independent list of length $n = \dim V$ is a basis (2.38).

Problem 6B.20. Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $\mathcal{E} \subseteq \mathcal{L}(V)$ is such that

$$ST = TS$$

for all $S, T \in \mathcal{E}$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.
This exercise strengthens Exercise 9(b) in Section 5E (in the context of inner product spaces) by asserting that the basis in that exercise can be chosen to be orthonormal.

Solution. Gram–Schmidt the basis supplied by Exercise 9(b) in Section 5E: the resulting orthonormal basis has the same partial spans, hence still triangularizes all of $\mathcal{E}$. The case $V = \{0\}$ is trivial, so set $n = \dim V \geq 1$. That exercise, applied to the nonzero finite-dimensional complex space V and the commuting family $\mathcal{E}$, gives a basis $v_1, \ldots, v_n$ of V with respect to which every element of $\mathcal{E}$ is upper triangular; equivalently, by 5.39,

$$\operatorname{span}(v_1, \ldots, v_k) \text{ is invariant under } T$$

for every $T \in \mathcal{E}$ and every $k \in \{1, \ldots, n\}$.
Applying the Gram–Schmidt procedure (6.32) to $v_1, \ldots, v_n$ produces an orthonormal list $e_1, \ldots, e_n$ with $\operatorname{span}(e_1, \ldots, e_k) = \operatorname{span}(v_1, \ldots, v_k)$ for each k; having length $n = \dim V$, it is an orthonormal basis of V (6.28). So each $\operatorname{span}(e_1, \ldots, e_k)$ is invariant under every $T \in \mathcal{E}$, and 5.39 in the other direction makes every $T \in \mathcal{E}$ upper triangular with respect to $e_1, \ldots, e_n$.

Problem 6B.21. Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, $T \in \mathcal{L}(V)$, and all eigenvalues of T have absolute value less than 1. Let $\epsilon > 0$. Prove that there exists a positive integer m such that $\|T^m v\| \leq \epsilon \|v\|$ for every $v \in V$.

Solution. Rescale a Schur basis so that T becomes a strict contraction for a new inner product, then transfer back by Exercise 16 of this section. The case $V = \{0\}$ is trivial, so let $n = \dim V \geq 1$. By Schur's theorem (6.38), applicable since $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional, there is an orthonormal basis $e_1, \ldots, e_n$ with $A = \mathcal{M}(T, (e_1, \ldots, e_n))$ upper triangular; by 5.41 its diagonal entries are the eigenvalues of T, so

$$r = \max\{|A_{1,1}|, \dots, |A_{n,n}|\} < 1.$$

Let $M = \max\{|A_{j,k}| : j < k\}$ (with $M = 0$ if $n = 1$), and put

$$t = \min\left\{1,\ \frac{1-r}{2n(M+1)}\right\} \in (0,1], \qquad f_k = t^k e_k.$$

Each f_k is a nonzero multiple of e_k, so $f_1, \dots, f_n$ is a basis, and substituting $e_j = t^{-j} f_j$ into $Te_k = \sum_{j\le k} A_{j,k} e_j$ gives $Tf_k = \sum_{j\le k} A_{j,k} t^{k-j} f_j$. Thus $B = \mathcal{M}(T, (f_1, \dots, f_n))$ is upper triangular with $B_{k,k} = A_{k,k}$ and $|B_{j,k}| \le Mt$ for $j < k$ (since $k - j \ge 1$ and $t \le 1$). Let $\langle \cdot, \cdot \rangle'$ be the inner product on V making $f_1, \dots, f_n$ orthonormal, obtained by transporting the standard inner product of $\mathbf{C}^n$ through the isomorphism $\sum a_k f_k \mapsto (a_1, \dots, a_n)$, and write $\|\cdot\|'$ for its norm.
Write $T = S + N$, where S and N have matrices with respect to $f_1, \dots, f_n$ equal to the diagonal and the strictly upper-triangular parts of B. For $v = \sum_k a_k f_k$ we get $Sv = \sum_k A_{k,k} a_k f_k$, so $\|Sv\|' \le r\|v\|'$, while the j-th coordinate of Nv is $\sum_{k>j} B_{j,k} a_k$, of modulus at most $Mt \sum_k |a_k| \le Mt\sqrt{n}\,\|v\|'$ by Cauchy–Schwarz (6.14); squaring and summing over the n values of j gives $\|Nv\|' \le nMt\|v\|'$. Since $nMt \le \frac{1-r}{2} \cdot \frac{M}{M+1} \le \frac{1-r}{2}$, the triangle inequality (6.17) yields

$$\|Tv\|' \le \|Sv\|' + \|Nv\|' \le \rho\|v\|', \qquad \rho = \tfrac{1+r}{2},$$

where $\rho \in [\frac{1}{2}, 1)$ because $0 \le r < 1$; iterating gives $\|T^m v\|' \le \rho^m \|v\|'$ for every positive integer m.
Finally, Exercise 16 of this section, applied to the two inner products on the finite-dimensional space V in each order, supplies $c_1, c_2 > 0$ with $\|w\| \le c_1\|w\|'$ and $\|w\|' \le c_2\|w\|$ for all $w \in V$, whence

$$\|T^m v\| \le c_1\|T^m v\|' \le c_1 \rho^m \|v\|' \le c_1 c_2 \rho^m \|v\|.$$

As $0 < \rho < 1$ we have $\rho^m \to 0$, so any m with $c_1 c_2 \rho^m \le \epsilon$ satisfies $\|T^m v\| \le \epsilon\|v\|$ for every $v \in V$.

Problem 6B.22. Suppose $C[-1,1]$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f, g \rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1,1]$. Let φ be the linear functional on $C[-1,1]$ defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in C[-1,1]$ such that

$$\varphi(f) = \langle f, g \rangle$$

for every $f \in C[-1,1]$.
This exercise shows that the Riesz representation theorem (6.42) does not hold on infinite-dimensional vector spaces without additional hypotheses on V and φ.

Solution. Test a putative g against $f(x) = x^2 g(x)$. Suppose $g \in C[-1,1]$ satisfies $f(0) = \int_{-1}^{1} fg$ for every $f \in C[-1,1]$. This f is continuous with $f(0) = 0$, so

$$0 = f(0) = \int_{-1}^{1} fg = \int_{-1}^{1} x^2 g(x)^2\, dx.$$

The integrand is continuous and nonnegative, and a continuous nonnegative h on $[-1,1]$ with $\int_{-1}^{1} h = 0$ vanishes identically: if $h(c) > 0$, continuity gives a subinterval $[\alpha, \beta] \subseteq [-1,1]$ of positive length on which $h > \frac{1}{2}h(c)$, whence $\int_{-1}^{1} h \ge \frac{1}{2}h(c)(\beta - \alpha) > 0$. So $x^2 g(x)^2 = 0$ throughout $[-1,1]$, giving $g(x) = 0$ for $x \ne 0$ and then $g(0) = \lim_{x\to 0} g(x) = 0$ by continuity.
Thus $g = 0$, so $f(0) = \langle f, 0 \rangle = 0$ for every $f \in C[-1,1]$ – contradicted by the constant function $f = 1$. Hence no such g exists.

Problem 6B.23. For all $u, v \in V$, define $d(u, v) = \|u - v\|$.

- (a) Show that d is a metric on V.
- (b) Show that if V is finite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).
- (c) Show that every finite-dimensional subspace of V is a closed subset of V (with respect to the metric d).

This exercise requires familiarity with metric spaces.

Solution. (a) The three axioms come straight from the norm. Nonnegativity holds since $\langle w, w\rangle \geq 0$; $d(u, v) = 0$ forces $u - v = 0$ by 6.9(a), and conversely $\|0\| = 0$; symmetry is 6.9(b) with $\lambda = -1$; and with $u - w = (u - v) + (v - w)$ the triangle inequality for the norm (6.17) gives

$$d(u, w) \leq \|u - v\| + \|v - w\| = d(u, v) + d(v, w).$$

(b) Coordinates with respect to an orthonormal basis turn a Cauchy sequence in V into n Cauchy sequences of scalars. The case $V = \{0\}$ is trivial, so let $e_1, \ldots, e_n$ be an orthonormal basis of V (6.35) and let (v_m) be Cauchy, with $a_{k,m} = \langle v_m, e_k\rangle$. Cauchy–Schwarz (6.14) and $\|e_k\| = 1$ give

$$|a_{k,m} - a_{k,j}| = |\langle v_m - v_j, e_k\rangle| \leq d(v_m, v_j),$$

so each $(a_{k,m})_m$ is Cauchy in $\mathbf{F}$, hence converges to some a_k by completeness of $\mathbb{R}$ and $\mathbb{C}$. Put $v = a_1 e_1 + \cdots + a_n e_n$, so $\langle v_m - v, e_k\rangle = a_{k,m} - a_k$ and 6.30(b) gives

$$d(v_m, v)^2 = \|v_m - v\|^2 = \sum_{k=1}^{n} |a_{k,m} - a_k|^2 \longrightarrow 0$$

as $m \to \infty$, the sum being one of finitely many terms each tending to 0. Thus $v_m \to v$ and d is complete.

(c) A finite-dimensional subspace U is complete, and a complete subspace of a metric space is closed. Precisely, U inherits the inner product of V, hence the metric d restricted to U, so (U, d) is complete by (b). If v lies in the closure of U, pick $u_m \in U$ with $d(u_m, v) < 1/m$; then (u_m) is Cauchy by the triangle inequality, so $u_m \to u$ for some $u \in U$, and

$$d(u, v) \leq d(u, u_m) + d(u_m, v) \longrightarrow 0,$$

forcing $d(u, v) = 0$ and hence $v = u \in U$ by (a). So U is closed in V.

6.3 Exercises 6C

Problem 6C.1. Suppose $v_1, \ldots, v_m \in V$. Prove that

$$\{v_1, \ldots, v_m\}^{\perp} = \big(\operatorname{span}(v_1, \ldots, v_m)\big)^{\perp}.$$

Solution. Both inclusions are one line. Write $U = \operatorname{span}(v_1, \ldots, v_m)$. Since $\{v_1, \ldots, v_m\} \subseteq U$, 6.48(e) gives $U^{\perp} \subseteq \{v_1, \ldots, v_m\}^{\perp}$. Conversely, if $\langle v_k, w\rangle = 0$ for each k and $u = a_1 v_1 + \cdots + a_m v_m \in U$, then additivity and homogeneity in the first slot give

$$\langle u, w\rangle = a_1\langle v_1, w\rangle + \cdots + a_m\langle v_m, w\rangle = 0,$$

so $w \in U^{\perp}$. Hence the two sets are equal.

Problem 6C.2. Suppose U is a subspace of V with basis $u_1, \ldots, u_m$ and

$$u_1, \ldots, u_m, v_1, \ldots, v_n$$

is a basis of V. Prove that if the Gram–Schmidt procedure is applied to the basis of V above, producing a list $e_1, \ldots, e_m, f_1, \ldots, f_n$, then $e_1, \ldots, e_m$ is an orthonormal basis of U and $f_1, \ldots, f_n$ is an orthonormal basis of $U^\perp$.

Solution. Gram–Schmidt preserves partial spans, which pins down both halves of the list. Here $\dim V = m + n$ and $\dim U = m$. Writing the given basis as $w_1, \ldots, w_{m+n}$ and the output as $g_1, \ldots, g_{m+n}$ (so $g_k = e_k$ for $k \leq m$ and $g_{m+j} = f_j$), 6.32 gives $\operatorname{span}(g_1, \ldots, g_k) = \operatorname{span}(w_1, \ldots, w_k)$ for every k; at $k = m$ this reads

$$\operatorname{span}(e_1, \ldots, e_m) = \operatorname{span}(u_1, \ldots, u_m) = U.$$

So $e_1, \ldots, e_m$ is an orthonormal spanning list of U, hence (being linearly independent by 6.25) an orthonormal basis of U.
Orthonormality of $g_1, \ldots, g_{m+n}$ gives $\langle e_k, f_j \rangle = 0$ for all $k \leq m$, so by Exercise 6C.1 and the displayed span,

$$\begin{aligned} f_j \in \{e_1, \ldots, e_m\}^\perp &= \big(\operatorname{span}(e_1, \ldots, e_m)\big)^\perp \\ &= U^\perp. \end{aligned}$$

Thus $f_1, \ldots, f_n$ is an orthonormal list in $U^\perp$, whose dimension is $\dim V - \dim U = n$ by 6.51; a list of that length is an orthonormal basis of $U^\perp$ by 6.28.

Problem 6C.3. Suppose U is the subspace of $\mathbf{R}^4$ defined by

$$U = \operatorname{span}\big((1, 2, 3, -4),\ (-5, 4, 3, 2)\big).$$

Find an orthonormal basis of U and an orthonormal basis of $U^\perp$.

Solution. The answers are

$$\tfrac{1}{\sqrt{30}}(1, 2, 3, -4), \qquad \tfrac{1}{\sqrt{12030}}(-77, 56, 39, 38)$$

for U, and

$$\tfrac{1}{\sqrt{139}}(3, 9, -7, 0), \qquad \tfrac{1}{\sqrt{55739}}(151, 36, 111, 139)$$

for $U^\perp$. Here is the computation, with $v_1 = (1, 2, 3, -4)$ and $v_2 = (-5, 4, 3, 2)$; neither is a multiple of the other, so $\dim U = 2$ and $\dim U^\perp = 2$ by 6.51.
For U, Gram–Schmidt (6.32) uses $\|v_1\|^2 = 30$ and $\langle v_2, v_1 \rangle = 4$:

$$\begin{aligned} v_2 - \tfrac{4}{30} v_1 &= \tfrac{1}{15}(-77, 56, 39, 38), \\ \|(-77, 56, 39, 38)\|^2 &= 12030, \end{aligned}$$

and scaling by a positive constant is harmless, giving the first display. For $U^\perp$, Exercise 6C.1 turns membership into the two equations $\langle v_1, x \rangle = \langle v_2, x \rangle = 0$, that is,

$$x_1 + 2x_2 + 3x_3 - 4x_4 = 0, \qquad -5x_1 + 4x_2 + 3x_3 + 2x_4 = 0.$$

Subtracting gives $x_2 = 3x_1 - 3x_4$, and then the first equation gives $3x_3 = 10x_4 - 7x_1$; writing $x_1 = 3a$, $x_4 = 3b$ yields

$$x = a(3, 9, -7, 0) + b(0, -9, 10, 3).$$

So $U^\perp = \operatorname{span}(w_1, w_2)$ for these two independent vectors. Gram–Schmidt again, with $\|w_1\|^2 = 139$ and $\langle w_2, w_1 \rangle = -151$:

$$\begin{aligned} w_2 + \tfrac{151}{139} w_1 &= \tfrac{3}{139}(151, 36, 111, 139), \\ \|(151, 36, 111, 139)\|^2 &= 55739. \end{aligned}$$

Each pair is orthonormal of length 2 inside a 2-dimensional space, hence a basis of it by 6.28. (Check!)

Problem 6C.4. Suppose $e_1, \dots, e_n$ is a list of vectors in V with $\|e_k\| = 1$ for each $k = 1, \dots, n$ and

$$\|v\|^2 = \big|\langle v, e_1\rangle\big|^2 + \cdots + \big|\langle v, e_n\rangle\big|^2$$

for all $v \in V$. Prove that $e_1, \dots, e_n$ is an orthonormal basis of V.
[This exercise provides a converse to 6.30(b).]

Solution. Orthogonality falls out of the hypothesis at $v = e_j$, and spanning from $U^\perp = \{0\}$. Fix j and take $v = e_j$:

$$\begin{aligned} 1 = \|e_j\|^2 &= \sum_{k=1}^{n} |\langle e_j, e_k\rangle|^2 \\ &= 1 + \sum_{k \neq j} |\langle e_j, e_k\rangle|^2, \end{aligned}$$

using $\langle e_j, e_j\rangle = \|e_j\|^2 = 1$. Every term of the remaining nonnegative sum is 0, so $\langle e_j, e_k\rangle = 0$ for $k \neq j$ and the list is orthonormal.
Let $U = \operatorname{span}(e_1, \dots, e_n)$, a finite-dimensional subspace of V. If $v \in U^\perp$ then $\langle e_k, v\rangle = 0$ for each k by Exercise 6C.1, hence $\langle v, e_k\rangle = 0$ by conjugate symmetry, and the hypothesis gives

$$\|v\|^2 = \sum_{k=1}^{n} |\langle v, e_k\rangle|^2 = 0,$$

so $v = 0$. Thus $U^\perp = \{0\}$, and 6.54 (applicable as U is finite-dimensional) gives $U = V$. An orthonormal spanning list is linearly independent (6.25), so $e_1, \dots, e_n$ is an orthonormal basis of V.

Problem 6C.5. Suppose that V is finite-dimensional and U is a subspace of V. Show that $P_{U^\perp} = I - P_U$, where I is the identity operator on V.

Solution. The decomposition $v = u + w$ works for both projections at once, read in the two orders. Both U and $U^\perp$ are finite-dimensional (2.25), so P_U and $P_{U^\perp}$ are defined (6.55). Given $v \in V$, write $v = u + w$ with $u \in U$ and $w \in U^\perp$ (6.49), so $P_U v = u$. Since $(U^\perp)^\perp = U$ by 6.52 (U is finite-dimensional), the rewriting $v = w + u$ is the decomposition of v along $U^\perp \oplus (U^\perp)^\perp$, whence

$$P_{U^\perp} v = w = v - u = v - P_U v = (I - P_U)v.$$

As v was arbitrary, $P_{U^\perp} = I - P_U$.

Problem 6C.6. Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$T = TP_{(\operatorname{null} T)^\perp} = P_{\operatorname{range} T}\, T.$$

Solution. Both identities say a projection changes nothing that T sees. Since V is finite-dimensional, so are $\operatorname{null} T$, $(\operatorname{null} T)^\perp$, and $\operatorname{range} T$ (3.21), so both projections are defined.
For the first, let $v \in V$ and write $v = u + x$ with $u \in \operatorname{null} T$ and $x \in (\operatorname{null} T)^\perp$ (6.49); by Exercise 6C.5 with $U = \operatorname{null} T$, or equivalently by 6.52, this gives $P_{(\operatorname{null} T)^\perp} v = x$. Since $Tu = 0$,

$$T\big(P_{(\operatorname{null} T)^\perp} v\big) = Tx = T(x + u) = Tv.$$

For the second, $Tv \in \operatorname{range} T$, so $P_{\operatorname{range} T}(Tv) = Tv$ by 6.57(b). As v was arbitrary, $T = TP_{(\operatorname{null} T)^\perp} = P_{\operatorname{range} T} T$.

Problem 6C.7. Suppose that X and Y are finite-dimensional subspaces of V. Prove that $P_X P_Y = 0$ if and only if $\langle x, y\rangle = 0$ for all $x \in X$ and all $y \in Y$.

Solution. Both directions run through null $P_X = X^{\perp}$ [6.57(e)], the orthogonality condition being exactly $Y \subseteq X^{\perp}$. Here P_X, P_Y are defined since X, Y are finite-dimensional (6.55).
(i) If $\langle x, y\rangle = 0$ for all $x \in X$, $y \in Y$, then for every $v \in V$ we have $P_Y v \in \operatorname{range} P_Y = Y \subseteq X^{\perp} = \operatorname{null} P_X$ by 6.57(d), (e), so $(P_X P_Y)v = 0$ and $P_X P_Y = 0$.
(ii) If $P_X P_Y = 0$, then for $y \in Y$ we have $P_Y y = y$ by 6.57(b), so

$$P_X y = P_X(P_Y y) = (P_X P_Y)y = 0,$$

putting $y \in \operatorname{null} P_X = X^{\perp}$; that is, $\langle x, y\rangle = 0$ for all $x \in X$, $y \in Y$.

Problem 6C.8. Suppose U is a finite-dimensional subspace of V and $v \in V$. Define a linear functional $\varphi\colon U \to \mathbf{F}$ by

$$\varphi(u) = \langle u, v\rangle$$

for all $u \in U$. By the Riesz representation theorem (6.42) as applied to the inner product space U, there exists a unique vector $w \in U$ such that

$$\varphi(u) = \langle u, w\rangle$$

for all $u \in U$. Show that $w = P_U v$.

Solution. The vector $v - w$ is orthogonal to U, which identifies w as the U-component of v. Indeed, for every $u \in U$, using additivity and conjugate homogeneity in the second slot [6.6(d), 6.6(e)],

$$\begin{aligned}\langle u, v - w\rangle &= \langle u, v\rangle - \langle u, w\rangle \\ &= \varphi(u) - \varphi(u) = 0,\end{aligned}$$

so $v - w \in U^{\perp}$. Thus $v = w + (v - w)$ with $w \in U$ and $v - w \in U^{\perp}$, which is the unique such decomposition since $V = U \oplus U^{\perp}$ (6.49, applicable as U is finite-dimensional). By the definition 6.55 of the orthogonal projection, $P_U v = w$.

Problem 6C.9. Suppose V is finite-dimensional. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in null P is orthogonal to every vector in range P. Prove that there exists a subspace U of V such that $P = P_U$.

Solution. Take $U = \operatorname{range} P$, a finite-dimensional subspace of V, so P_U is defined (6.55). For $v \in V$, write $v = Pv + (v - Pv)$. The first term lies in U, while

$$P(v - Pv) = Pv - P^2 v = 0,$$

so $v - Pv \in \operatorname{null} P$, which by hypothesis is orthogonal to range $P = U$; hence $v - Pv \in U^{\perp}$. As $V = U \oplus U^{\perp}$ (6.49), this decomposition is the unique one, so $P_U v = Pv$ by 6.55. Since v was arbitrary, $P = P_U$.

Problem 6C.10. Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and

$$\|Pv\| \leq \|v\|$$

for every $v \in V$. Prove that there exists a subspace U of V such that $P = P_U$.

Solution. By Exercise 6C.9 it suffices to show $\operatorname{range} P \perp \operatorname{null} P$; then $P = P_U$ with $U = \operatorname{range} P$. So let $u \in \operatorname{range} P$ and $0 \neq w \in \operatorname{null} P$ (the case $w = 0$ being trivial by 6.6(c)). Writing $u = Px$ gives $Pu = P^2x = Px = u$, so $P(u + \lambda w) = u$ for every $\lambda \in \mathbf{F}$ and the hypothesis applied to $v = u + \lambda w$ yields $\|u\| \leq \|u + \lambda w\|$. Expanding,

$$\begin{aligned}\|u + \lambda w\|^2 = \|u\|^2 + |\lambda|^2\|w\|^2 \\ + 2\operatorname{Re}\big(\overline{\lambda}\langle u, w\rangle\big).\end{aligned}$$

Now take $\lambda = -\langle u, w\rangle/\|w\|^2$, legitimate since $w \neq 0$, so that $\overline{\lambda}\langle u, w\rangle = -|\langle u, w\rangle|^2/\|w\|^2$ is real and $|\lambda|^2\|w\|^2 = |\langle u, w\rangle|^2/\|w\|^2$. The expansion becomes

$$\|u\|^2 \leq \|u + \lambda w\|^2 = \|u\|^2 - \frac{|\langle u, w\rangle|^2}{\|w\|^2},$$

forcing the nonnegative quantity $|\langle u, w\rangle|^2/\|w\|^2$ to vanish. Hence $\langle u, w\rangle = 0$, and Exercise 6C.9 gives $P = P_{\operatorname{range} P}$.

Problem 6C.11. Suppose $T \in \mathcal{L}(V)$ and U is a finite-dimensional subspace of V. Prove that

$$U \text{ is invariant under } T \iff P_U T P_U = T P_U.$$

Solution. Both directions are 6.57(b) and 6.57(d) applied once each; P_U is defined since U is finite-dimensional (6.55).
(i) If U is invariant under T, then for $v \in V$ we have $P_U v \in \operatorname{range} P_U = U$ [6.57(d)], so $TP_U v \in U$ and P_U fixes it [6.57(b)]: $P_U T P_U v = T P_U v$.
(ii) If $P_U T P_U = T P_U$, then for $u \in U$ we have $P_U u = u$ [6.57(b)], so

$$Tu = TP_U u = P_U T P_U u = P_U(Tu) \in \operatorname{range} P_U = U,$$

so U is invariant under T.

Problem 6C.12. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V. Prove that

$$U \text{ and } U^\perp \text{ are both invariant under } T \iff P_U T = T P_U.$$

Solution. Both directions come from $V = U \oplus U^\perp$ (6.49) together with $\operatorname{range} P_U = U$ and $\operatorname{null} P_U = U^\perp$ [6.57(d), (e)]; U is finite-dimensional since V is.
(i) Suppose U and $U^\perp$ are invariant under T. Writing $v = u + w$ with $u \in U$, $w \in U^\perp$, we get $TP_U v = Tu$, while $Tv = Tu + Tw$ with $Tu \in U$ and $Tw \in U^\perp$ is the decomposition defining $P_U T v$, so $P_U T v = Tu = T P_U v$.
(ii) Suppose $P_U T = T P_U$. For $u \in U$, $P_U u = u$ [6.57(b)] gives $Tu = TP_U u = P_U T u \in U$; for $w \in U^\perp = \operatorname{null} P_U$,

$$P_U(Tw) = TP_U w = T0 = 0,$$

so $Tw \in \operatorname{null} P_U = U^\perp$. Thus both U and $U^\perp$ are invariant under T.

Problem 6C.13. Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. For each $v \in V$, let φ_v denote the linear functional on V defined by

$$\varphi_v(u) = \langle u, v\rangle$$

for all $u \in V$.
(a) Show that $v \mapsto \varphi_v$ is an injective linear map from V to V'.
(b) Use (a) and a dimension-counting argument to show that $v \mapsto \varphi_v$ is an isomorphism from V onto V'.

[The purpose of this exercise is to give an alternative proof of the Riesz representation theorem (6.42 and 6.58) when $\mathbf{F} = \mathbf{R}$. Thus you should not use the Riesz representation theorem as a tool in your solution.]

Solution. (a) Write $\Phi(v) = \varphi_v$; each φ_v lies in V' by 6.6(a). Linearity of Φ is additivity and homogeneity in the second slot [6.6(d), 6.6(e)]: for all $u \in V$,

$$\varphi_{v_1+v_2}(u) = \langle u, v_1\rangle + \langle u, v_2\rangle, \qquad \varphi_{\lambda v}(u) = \overline{\lambda}\langle u, v\rangle = \lambda\varphi_v(u),$$

the last step using $\overline{\lambda} = \lambda$ as $\mathbf{F} = \mathbf{R}$. For injectivity, $\varphi_v = 0$ evaluated at $u = v$ gives $\|v\|^2 = \langle v, v\rangle = 0$, so $v = 0$ and $\operatorname{null}\Phi = \{0\}$ (3.15).
(b) By 3.111, $\dim V' = \dim V < \infty$, so the fundamental theorem of linear maps (3.21) applied to Φ with $\dim\operatorname{null}\Phi = 0$ gives

$$\dim\operatorname{range}\Phi = \dim V = \dim V'.$$

A subspace of V' of full dimension is V' (2.39), so Φ is surjective, hence an isomorphism.

Problem 6C.14. Suppose that $e_1, \dots, e_n$ is an orthonormal basis of V. Explain why the dual basis (see 3.112) of $e_1, \dots, e_n$ is $e_1, \dots, e_n$ under the identification of V' with V provided by the Riesz representation theorem (6.58).

Solution. Because φ_{e_j} and the j-th dual-basis functional φ_j of 3.112 agree on the basis $e_1, \dots, e_n$: orthonormality gives

$$\varphi_{e_j}(e_k) = \langle e_k, e_j\rangle = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j, \end{cases}$$

which is exactly the defining property of φ_j. Two linear maps agreeing on a basis are equal, so $\varphi_j = \varphi_{e_j}$ for each j; that is, under the correspondence $v \leftrightarrow \varphi_v$ of 6.58 the dual basis of $e_1, \dots, e_n$ is $e_1, \dots, e_n$.

Problem 6C.15. In $\mathbb{R}^4$, let

$$U = \operatorname{span}\big((1,1,0,0), (1,1,1,2)\big).$$

Find $u \in U$ such that $\|u - (1,2,3,4)\|$ is as small as possible.

Solution. The answer is $u = \left(\frac{3}{2}, \frac{3}{2}, \frac{11}{5}, \frac{22}{5}\right)$, since by 6.61 the minimizer is $u = P_U v$ with $v = (1,2,3,4)$. Gram–Schmidt (6.32) on the basis $(1,1,0,0), (1,1,1,2)$ of U gives

$$e_1 = \tfrac{1}{\sqrt{2}}(1,1,0,0), \qquad e_2 = \tfrac{1}{\sqrt{5}}(0,0,1,2),$$

because $(1,1,1,2) - \sqrt{2}\,e_1 = (0,0,1,2)$, of norm $\sqrt{5}$. Then $\langle v, e_1\rangle = 3/\sqrt{2}$ and $\langle v, e_2\rangle = 11/\sqrt{5}$, so 6.57(i) gives

$$\begin{aligned} P_U v &= \tfrac{3}{2}(1,1,0,0) + \tfrac{11}{5}(0,0,1,2) \\ &= \left(\tfrac{3}{2}, \tfrac{3}{2}, \tfrac{11}{5}, \tfrac{22}{5}\right). \end{aligned}$$

The minimum distance is $\|v - u\| = \sqrt{13/10}$.

Problem 6C.16. Suppose $C[-1,1]$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f, g\rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1,1]$. Let U be the subspace of $C[-1,1]$ defined by

$$U = \{f \in C[-1,1] : f(0) = 0\}.$$

(a) Show that $U^{\perp} = \{0\}$.
(b) Show that 6.49 and 6.52 do not hold without the finite-dimensional hypothesis.

Solution. (a) Test $g \in U^{\perp}$ against $f(x) = x^2 g(x)$, which is continuous with $f(0) = 0$, hence in U:

$$0 = \langle f, g\rangle = \int_{-1}^{1} x^2 g(x)^2 \, dx.$$

The integrand is continuous and nonnegative, and such a function with zero integral vanishes identically (if $h(x_0) > 0$, continuity gives a nondegenerate subinterval on which $h > \frac{1}{2}h(x_0)$, forcing $\int_{-1}^{1} h > 0$). So $g(x) = 0$ for $x \neq 0$, and $g(0) = \lim_{x\to 0} g(x) = 0$ by continuity. Hence $U^{\perp} = \{0\}$.
(b) This U is infinite-dimensional (it contains $x, x^2, x^3, \ldots$, polynomials of distinct degrees), and the constant function 1 lies in $C[-1,1] \setminus U$. By (a),

$$\begin{aligned} U + U^{\perp} &= U \subsetneq C[-1,1], \\ (U^{\perp})^{\perp} &= \{0\}^{\perp} = C[-1,1] \neq U, \end{aligned}$$

so the conclusions of 6.49 and of 6.52 both fail for this U.

Problem 6C.17. Find $p \in \mathcal{P}_3(\mathbb{R})$ such that $p(0) = 0$, $p'(0) = 0$, and

$$\int_0^1 \left|2 + 3x - p(x)\right|^2 dx$$

is as small as possible.

Solution. The answer is $p(x) = 24x^2 - \frac{203}{10}x^3$. Work in $C[0,1]$ with $\langle f, g\rangle = \int_0^1 fg$ and set $v(x) = 2+3x$. The conditions $p(0) = p'(0) = 0$ kill the constant and linear coefficients, so the admissible set is

$$U = \operatorname{span}(x^2, x^3),$$

and by 6.61 the minimizer of $\|v - p\|^2$ over $p \in U$ is $p = P_U v$, characterized by $p \in U$ with $v - p \in U^{\perp}$, which by Exercise 6C.1 means $\langle v - p, x^2\rangle = \langle v - p, x^3\rangle = 0$. Since

$$\langle v, x^2\rangle = \tfrac{2}{3} + \tfrac{3}{4} = \tfrac{17}{12}, \qquad \langle v, x^3\rangle = \tfrac{1}{2} + \tfrac{3}{5} = \tfrac{11}{10},$$

writing $p = ax^2 + bx^3$ gives the normal equations

$$\begin{aligned} \tfrac{a}{5} + \tfrac{b}{6} &= \tfrac{17}{12}, \\ \tfrac{a}{6} + \tfrac{b}{7} &= \tfrac{11}{10}, \end{aligned}$$

whose solution is $a = 24$, $b = -\frac{203}{10}$. (Check!)

Problem 6C.18. Find $p \in \mathcal{P}_5(\mathbb{R})$ that makes

$$\int_{-\pi}^{\pi} \left|\sin x - p(x)\right|^2 dx$$

as small as possible.
The polynomial 6.65 is an excellent approximation to the answer to this exercise, but here you are asked to find the exact solution, which involves powers of π. A computer that can perform symbolic integration should help.

Solution. The answer is

$$\begin{aligned} p(x) = {} & \frac{105(\pi^4 - 153\pi^2 + 1485)}{8\pi^6}\, x \\ & - \frac{315(\pi^4 - 125\pi^2 + 1155)}{4\pi^8}\, x^3 \\ & + \frac{693(\pi^4 - 105\pi^2 + 945)}{8\pi^{10}}\, x^5. \end{aligned}$$

Work in $C[-\pi,\pi]$ with the inner product 6.64, let $v(x) = \sin x$, and let $U = \mathcal{P}_5(\mathbf{R})$. By 6.61 the minimizer is $p = P_U v$, characterized by $p \in U$ with $v - p \in U^\perp$ [6.49 and 6.57(f)], which by Exercise 6C.1 applied to the spanning list $1, x, \ldots, x^5$ reads

$$\int_{-\pi}^{\pi} x^k \big(\sin x - p(x)\big)\,dx = 0, \qquad k = 0, 1, \ldots, 5.$$

Parity supplies the integrals: $\int_{-\pi}^{\pi} x^m dx$ is 0 for m odd and $2\pi^{m+1}/(m+1)$ for m even, while $\int_{-\pi}^{\pi} x^k \sin x\, dx$ is 0 for k even and $2I_k$ for k odd, where $I_n = \int_0^{\pi} x^n \sin x\, dx$ satisfies $I_n = \pi^n - n(n-1)I_{n-2}$ by parts twice, with $I_1 = \pi$, hence

$$I_1 = \pi, \qquad I_3 = \pi^3 - 6\pi, \qquad I_5 = \pi^5 - 20\pi^3 + 120\pi.$$

Write $p(x) = a_0 + a_1 x + \cdots + a_5 x^5$. For k even the equation collapses to $\int_{-\pi}^{\pi} x^k q = 0$ with $q(x) = a_0 + a_2 x^2 + a_4 x^4$; thus q is orthogonal to $1, x^2, x^4$, hence to q itself, so $q = 0$ and $a_0 = a_2 = a_4 = 0$. For k odd, setting $a = a_1$, $b = a_3$, $c = a_5$ and dividing the $k = 1, 3, 5$ equations by $2\pi^3, 2\pi^5, 2\pi^7$, the substitution $t = \pi^2$, $A = a$, $B = \pi^2 b$, $C = \pi^4 c$ gives

$$\begin{aligned} \tfrac{A}{3} + \tfrac{B}{5} + \tfrac{C}{7} &= \tfrac{1}{t}, \\ \tfrac{A}{5} + \tfrac{B}{7} + \tfrac{C}{9} &= \tfrac{1}{t} - \tfrac{6}{t^2}, \\ \tfrac{A}{7} + \tfrac{B}{9} + \tfrac{C}{11} &= \tfrac{1}{t} - \tfrac{20}{t^2} + \tfrac{120}{t^3}. \end{aligned}$$

The coefficient matrix has determinant $\frac{256}{22920975} \neq 0$, and elimination gives

$$\begin{aligned} A &= \tfrac{105}{8t} - \tfrac{16065}{8t^2} + \tfrac{155925}{8t^3}, \\ B &= -\tfrac{315}{4t} + \tfrac{39375}{4t^2} - \tfrac{363825}{4t^3}, \\ C &= \tfrac{693}{8t} - \tfrac{72765}{8t^2} + \tfrac{654885}{8t^3}. \end{aligned}$$

(Check! – substitute and collect powers of $1/t$.) Undoing $a = A$, $b = B/\pi^2$, $c = C/\pi^4$, $t = \pi^2$ gives the displayed p, whose coefficients are $0.987862\ldots$, $-0.155271\ldots$, $0.00564312\ldots$, matching 6.65.

Problem 6C.19. Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is an orthogonal projection of V onto some subspace of V. Prove that $P^\dagger = P$.

Solution. Write $P = P_U$, where U is a (necessarily finite-dimensional) subspace of V. By 6.57(d) and 6.57(e),

$$\operatorname{range} P = U, \qquad (\operatorname{null} P)^\perp = (U^\perp)^\perp = U,$$

the last equality by 6.52, which applies since U is finite-dimensional. By 6.57(b), $P|_U$ is the identity map of U, hence so is its inverse $\big(P|_{(\operatorname{null} P)^\perp}\big)^{-1} \colon \operatorname{range} P \to (\operatorname{null} P)^\perp$. Thus 6.68 gives, for every $v \in V$,

$$P^\dagger v = \big(P|_{(\operatorname{null} P)^\perp}\big)^{-1} P_{\operatorname{range} P}\, v = \big(P|_U\big)^{-1}(Pv) = Pv,$$

using $P_{\operatorname{range} P} = P_U = P$ and $Pv \in U$. Hence $P^\dagger = P$.

Problem 6C.20. Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$\operatorname{null} T^\dagger = (\operatorname{range} T)^\perp \qquad \text{and} \qquad \operatorname{range} T^\dagger = (\operatorname{null} T)^\perp.$$

Solution. Set $S = T|_{(\operatorname{null} T)^\perp}$, which by 6.67 is an invertible map of $(\operatorname{null} T)^\perp$ onto $\operatorname{range} T$ (both finite-dimensional, as V is), so that 6.68 reads $T^\dagger w = S^{-1} P_{\operatorname{range} T} w$ for $w \in W$.
(i) Because S^{-1} is injective,

$$T^\dagger w = 0 \iff P_{\operatorname{range} T} w = 0 \iff w \in (\operatorname{range} T)^\perp,$$

the second equivalence by 6.57(e). Hence $\operatorname{null} T^\dagger = (\operatorname{range} T)^\perp$.
(ii) Every $T^\dagger w = S^{-1}(P_{\operatorname{range} T} w)$ lies in $\operatorname{range} S^{-1} = (\operatorname{null} T)^\perp$, giving $\subseteq$. Conversely, for $v \in (\operatorname{null} T)^\perp$ put $w = Tv = Sv \in \operatorname{range} T$; then $P_{\operatorname{range} T} w = w$ by 6.57(b), so

$$T^\dagger w = S^{-1} w = S^{-1}(Sv) = v.$$

Hence $\operatorname{range} T^\dagger = (\operatorname{null} T)^\perp$.

Problem 6C.21. Suppose $T \in \mathcal{L}(\mathbf{F}^3, \mathbf{F}^2)$ is defined by

$$T(a, b, c) = (a + b + c,\ 2b + 3c).$$

(a) For $(x, y) \in \mathbf{F}^2$, find a formula for $T^\dagger(x, y)$.
(b) Verify that the equation $TT^\dagger = P_{\operatorname{range} T}$ from 6.69(b) holds with the formula for $T^\dagger$ obtained in (a).
(c) Verify that the equation $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ from 6.69(c) holds with the formula for $T^\dagger$ obtained in (a).

Solution. (a) $T^\dagger(x, y) = \frac{1}{14}\big(13x - 5y,\ 3x + y,\ -2x + 4y\big)$.
Here $\mathbf{F}^n$ carries the Euclidean inner product. Since $T(1, 0, 0) = (1, 0)$ and $T(0, 1, 0) = (1, 2)$ are linearly independent, $\operatorname{range} T = \mathbf{F}^2$ and so $P_{\operatorname{range} T} = I$; solving $a + b + c = 0 = 2b + 3c$ gives

$$\operatorname{null} T = \operatorname{span}\big((1, -3, 2)\big), \qquad (\operatorname{null} T)^\perp = \{(a, b, c) : a - 3b + 2c = 0\},$$

the description of the orthogonal complement using that $(1, -3, 2)$ has real entries. By 6.67 and 6.68, $T^\dagger(x, y) = S^{-1}(x, y)$ with $S = T|_{(\operatorname{null} T)^\perp}$; that is, $T^\dagger(x, y)$ is the unique (a, b, c) with

$$a + b + c = x, \qquad 2b + 3c = y, \qquad a - 3b + 2c = 0.$$

Subtracting the third equation from the first gives $c = 4b - x$; then the second gives $14b = 3x + y$, and

$$b = \frac{3x + y}{14}, \qquad c = \frac{-2x + 4y}{14}, \qquad a = x - b - c = \frac{13x - 5y}{14},$$

which is the stated formula. (Check! the three equations.)
(b) By the three equations of (a), $TT^\dagger(x, y) = (x, y)$, so $TT^\dagger = I$; and $P_{\operatorname{range} T} = I$ since $\operatorname{range} T = \mathbf{F}^2$. Hence $TT^\dagger = P_{\operatorname{range} T}$.
(c) For $(a, b, c) \in \mathbf{F}^3$ put $x = a + b + c$ and $y = 2b + 3c$, so $T(a, b, c) = (x, y)$ and, by (a),

$$\begin{aligned} T^\dagger T(a, b, c) &= \tfrac{1}{14}\big(13x - 5y,\ 3x + y,\ -2x + 4y\big) \\ &= \tfrac{1}{14}\Big(13(a + b + c) - 5(2b + 3c),\ 3(a + b + c) + (2b + 3c),\ -2(a + b + c) + 4(2b + 3c)\Big) \\ &= \tfrac{1}{14}\big(13a + 3b - 2c,\ 3a + 5b + 6c,\ -2a + 6b + 10c\big). \end{aligned}$$

By Exercise 6C.5, $P_{(\operatorname{null} T)^\perp} = I - P_{\operatorname{null} T}$, and since $\|(1, -3, 2)\|^2 = 14$, formula 6.56 gives

$$P_{\operatorname{null} T}(a, b, c) = \frac{\langle (a, b, c), (1, -3, 2) \rangle}{14}(1, -3, 2) = \frac{a - 3b + 2c}{14}(1, -3, 2).$$

Hence

$$\begin{aligned}P_{(\operatorname{null}T)^{\perp}}(a,b,c) &= (a,b,c) - \frac{a-3b+2c}{14}(1,-3,2)\\ &= \tfrac{1}{14}\Big(14a-(a-3b+2c),\ 14b+3(a-3b+2c),\ 14c-2(a-3b+2c)\Big)\\ &= \tfrac{1}{14}\big(13a+3b-2c,\ 3a+5b+6c,\ -2a+6b+10c\big).\end{aligned}$$

The two displays agree for every (a,b,c), so $T^{\dagger}T = P_{(\operatorname{null}T)^{\perp}}$.

Problem 6C.22. Suppose V is finite-dimensional and $T \in \mathcal{L}(V,W)$. Prove that

$$TT^{\dagger}T = T \quad \text{and} \quad T^{\dagger}TT^{\dagger} = T^{\dagger}.$$

Both formulas above clearly hold if T is invertible because in that case we can replace $T^{\dagger}$ with T^{-1}.

Solution. Both identities come from 6.69 together with $TP_{(\operatorname{null}T)^{\perp}} = T$, which is Exercise 6C.6: writing $v = x + u$ with $x \in (\operatorname{null}T)^{\perp}$ and $u \in \operatorname{null}T$ (possible by 6.49, since $\operatorname{null}T$ is finite-dimensional, and this is the decomposition defining $P_{(\operatorname{null}T)^{\perp}}$ in 6.55 because $\big((\operatorname{null}T)^{\perp}\big)^{\perp} = \operatorname{null}T$ by 6.52), we get

$$TP_{(\operatorname{null}T)^{\perp}}v = Tx = Tx + Tu = Tv.$$

(i) By 6.69(c), $T^{\dagger}T = P_{(\operatorname{null}T)^{\perp}}$, so

$$TT^{\dagger}T = TP_{(\operatorname{null}T)^{\perp}} = T.$$

(ii) By 6.69(b), $TT^{\dagger} = P_{\operatorname{range}T}$, so for every $w \in W$ the definition 6.68 and the idempotence $P_{\operatorname{range}T}^2 = P_{\operatorname{range}T}$ (6.57(g)) give

$$\begin{aligned}T^{\dagger}TT^{\dagger}w &= \big(T|_{(\operatorname{null}T)^{\perp}}\big)^{-1}P_{\operatorname{range}T}^2\, w\\ &= \big(T|_{(\operatorname{null}T)^{\perp}}\big)^{-1}P_{\operatorname{range}T}\, w = T^{\dagger}w.\end{aligned}$$

Problem 6C.23. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V,W)$. Prove that

$$\big(T^{\dagger}\big)^{\dagger} = T.$$

The equation above is analogous to the equation $(T^{-1})^{-1} = T$ that holds if T is invertible.

Solution. Write $N = \operatorname{null}T$, $R = \operatorname{range}T$, and $S = T|_{N^{\perp}}$, which by 6.67 is invertible from $N^{\perp}$ onto R, so that 6.68 reads $T^{\dagger}w = S^{-1}P_R w$. By Exercise 6C.20,

$$\operatorname{null}T^{\dagger} = R^{\perp}, \qquad \operatorname{range}T^{\dagger} = N^{\perp},$$

and since R is finite-dimensional, 6.52 gives $\big(\operatorname{null}T^{\dagger}\big)^{\perp} = \big(R^{\perp}\big)^{\perp} = R$. So 6.68 applied to $T^{\dagger} \in \mathcal{L}(W,V)$ (legitimate because W is finite-dimensional) reads

$$\big(T^{\dagger}\big)^{\dagger}v = \big(T^{\dagger}|_R\big)^{-1}P_{N^{\perp}}\, v \qquad \text{for every } v \in V.$$

For $w \in R$ we have $P_R w = w$ by 6.57(b), so $T^{\dagger}|_R = S^{-1}$ as a map of R onto $N^{\perp}$, whence $\big(T^{\dagger}|_R\big)^{-1} = S$. Therefore

$$\big(T^{\dagger}\big)^{\dagger}v = S\big(P_{N^{\perp}}v\big) = T\big(P_{N^{\perp}}v\big) = Tv,$$

the second equality because $P_{N^{\perp}}v \in N^{\perp}$ and the third by Exercise 6C.6. Hence $\big(T^{\dagger}\big)^{\dagger} = T$.

7 Operators on Inner Product Spaces

Contents

7.1 Exercises 7A

Problem 7A.1. Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1}).$$

Find a formula for $T^*(z_1, \dots, z_n)$.

Solution. The adjoint of the forward shift is the backward shift:

$$T^*(z_1, \dots, z_n) = (z_2, \dots, z_n, 0).$$

Indeed, with $\mathbf{F}^n$ carrying the Euclidean inner product, for $z, w \in \mathbf{F}^n$

$$\langle Tz, w\rangle = \sum_{k=1}^{n-1} z_k \overline{w_{k+1}} = \big\langle z, (w_2, \dots, w_n, 0)\big\rangle,$$

the second equality because the omitted term $z_n\overline{0}$ vanishes. Comparing with the definition 7.1 of the adjoint gives $\langle z, T^*w - (w_2, \dots, w_n, 0)\rangle = 0$ for every z; taking z to be that second slot forces $T^*w = (w_2, \dots, w_n, 0)$.

Problem 7A.2. Suppose $T \in \mathcal{L}(V, W)$. Prove that

$$T = 0 \iff T^* = 0 \iff T^*T = 0 \iff TT^* = 0.$$

Solution. (i) $T = 0 \iff T^* = 0$: if $T = 0$ then $\|T^*w\|^2 = \langle T^*w, T^*w\rangle = \langle T(T^*w), w\rangle = 0$ for every $w \in W$ by 7.1, so $T^* = 0$; applying this to T^* and using $(T^*)^* = T$ (7.5(c)) gives the converse.
(ii) $T = 0 \iff T^*T = 0$: the forward direction is immediate, and conversely, for $v \in V$, 7.1 gives

$$\|Tv\|^2 = \langle Tv, Tv\rangle = \langle v, (T^*T)v\rangle = 0,$$

so $T = 0$.
(iii) $T^* = 0 \iff TT^* = 0$: this is (ii) applied to T^*, again using $(T^*)^* = T$.
Combining (i), (ii), (iii) gives the four stated equivalences.

Problem 7A.3. Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Prove that

$$\lambda \text{ is an eigenvalue of } T \iff \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

Solution. By 7.5(a), 7.5(b), and 7.5(e), $(T - \lambda I)^* = T^* - \overline{\lambda} I$. Hence, V being finite-dimensional,

$$\begin{aligned}
\lambda \text{ is an eigenvalue of } T &\iff T - \lambda I \text{ is not injective} \\
&\iff T - \lambda I \text{ is not surjective} \\
&\iff \operatorname{range}(T - \lambda I) \neq V \\
&\iff \big(\operatorname{range}(T - \lambda I)\big)^{\perp} \neq \{0\} \\
&\iff \operatorname{null}\big((T - \lambda I)^*\big) \neq \{0\} \\
&\iff \operatorname{null}\big(T^* - \overline{\lambda} I\big) \neq \{0\} \\
&\iff \overline{\lambda} \text{ is an eigenvalue of } T^*.
\end{aligned}$$

Here the second equivalence is 3.65 (applicable as $T - \lambda I$ is an operator on the finite-dimensional V), the fourth is the contrapositive of 6.54 applied to the subspace $\operatorname{range}(T - \lambda I)$, the fifth is 7.6(a), and the sixth is the identity above.

Problem 7A.4. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that

$$U \text{ is invariant under } T \iff U^\perp \text{ is invariant under } T^*.$$

Solution. ($\Longrightarrow$) Suppose U is invariant under T and $w \in U^\perp$. For every $u \in U$, the definition 7.1 of the adjoint gives

$$\langle u, T^*w\rangle = \langle Tu, w\rangle = 0,$$

since $Tu \in U$ and $w \in U^\perp$. Hence $T^*w \in U^\perp$, so $U^\perp$ is invariant under T^*.
($\Longleftarrow$) Apply the implication just proved to the subspace $U^\perp$ and the operator T^*: it gives that $(U^\perp)^\perp$ is invariant under $(T^*)^*$. Now $(U^\perp)^\perp = U$ by 6.52 (valid as U is finite-dimensional) and $(T^*)^* = T$ by 7.5(c), so U is invariant under T.

Problem 7A.5. Suppose $T \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Prove that

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^*f_1\|^2 + \cdots + \|T^*f_m\|^2.$$

[The numbers $\|Te_1\|^2, \dots, \|Te_n\|^2$ in the equation above depend on the orthonormal basis $e_1, \dots, e_n$, but the right side of the equation does not depend on $e_1, \dots, e_n$. Thus the equation above shows that the sum on the left side does not depend on which orthonormal basis $e_1, \dots, e_n$ is used.]

Solution. Both sides equal $\sum_{j,k} |\langle Te_j, f_k\rangle|^2$. Indeed, 6.30(b) applied in W and then in V gives

$$\sum_{j=1}^{n} \|Te_j\|^2 = \sum_{j=1}^{n}\sum_{k=1}^{m}\bigl|\langle Te_j, f_k\rangle\bigr|^2,$$
$$\sum_{k=1}^{m} \|T^*f_k\|^2 = \sum_{k=1}^{m}\sum_{j=1}^{n}\bigl|\langle T^*f_k, e_j\rangle\bigr|^2,$$

while the definition 7.1 of the adjoint and conjugate symmetry give

$$\langle T^*f_k, e_j\rangle = \overline{\langle e_j, T^*f_k\rangle} = \overline{\langle Te_j, f_k\rangle},$$

so the two double sums have identical terms. Hence the two sides are equal.

Problem 7A.6. Suppose $T \in \mathcal{L}(V, W)$. Prove that
(a) T is injective $\iff$ T^* is surjective;
(b) T is surjective $\iff$ T^* is injective.

Solution. We use twice, for a subspace U of a finite-dimensional inner product space X, that

$$U^\perp = X \iff U = \{0\} \qquad \text{and} \qquad U^\perp = \{0\} \iff U = X.$$

The second is 6.54; the first follows from $\dim U^\perp = \dim X - \dim U$ (6.51) together with 2.39.
(a) By 7.6(b), $\operatorname{range} T^* = (\operatorname{null} T)^\perp$, so

$$\begin{aligned} T^* \text{ is surjective} &\iff \operatorname{range} T^* = V \\ &\iff (\operatorname{null} T)^\perp = V \\ &\iff \operatorname{null} T = \{0\} \\ &\iff T \text{ is injective,} \end{aligned}$$

the third equivalence by the first displayed fact with $X = V$, $U = \operatorname{null} T$, and the fourth by 3.15.

(b) By 7.6(a), $\operatorname{null} T^* = (\operatorname{range} T)^\perp$, so

$$\begin{aligned} T^* \text{ is injective} &\iff \operatorname{null} T^* = \{0\} \\ &\iff (\operatorname{range} T)^\perp = \{0\} \\ &\iff \operatorname{range} T = W \\ &\iff T \text{ is surjective,} \end{aligned}$$

the first equivalence by 3.15 and the third by the second displayed fact with $X = W$, $U = \operatorname{range} T$.

Problem 7A.7. Prove that if $T \in \mathcal{L}(V, W)$, then
(a) $\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$;
(b) $\dim \operatorname{range} T^* = \dim \operatorname{range} T$.

Solution. Both parts come from 7.6 combined with 6.51 and the fundamental theorem of linear maps 3.21 (applicable since V and W are finite-dimensional).
(a) By 7.6(a) and 6.51 inside W, then 3.21 applied to T,

$$\begin{aligned} \dim \operatorname{null} T^* &= \dim(\operatorname{range} T)^\perp = \dim W - \dim \operatorname{range} T \\ &= \dim W - \big(\dim V - \dim \operatorname{null} T\big) \\ &= \dim \operatorname{null} T + \dim W - \dim V. \end{aligned}$$

(b) By 7.6(b) and 6.51 inside V, then 3.21 applied to T,

$$\dim \operatorname{range} T^* = \dim(\operatorname{null} T)^\perp = \dim V - \dim \operatorname{null} T = \dim \operatorname{range} T.$$

Problem 7A.8. Suppose A is an m-by-n matrix with entries in $\mathbf{F}$. Use (b) in Exercise 7 to prove that the row rank of A equals the column rank of A.
[This exercise asks for yet another alternative proof of a result that was previously proved in 3.57 and 3.133.]

Solution. Define $T\colon \mathbf{F}^{n,1} \to \mathbf{F}^{m,1}$ by $Tx = Ax$, where both spaces carry their standard inner products, for which the standard bases are orthonormal. Then $\mathcal{M}(T) = A$ with respect to those bases (the k^{th} column of A is $Ae_k = Te_k$), so 7.9 gives $\mathcal{M}(T^*) = A^*$, the conjugate transpose. Hence

$$\begin{aligned} \text{column rank of } A &= \text{column rank of } \mathcal{M}(T) \\ &= \dim \operatorname{range} T \\ &= \dim \operatorname{range} T^* \\ &= \text{column rank of } \mathcal{M}(T^*) \\ &= \text{column rank of } A^*, \end{aligned}$$

where the second and fourth equalities come from 3.78 and the third equality comes from (b) in Exercise 7.
Finally, the column rank of A^* equals the row rank of A: writing $r_1, \dots, r_m$ for the rows of A, the definition 7.7 makes the columns of A^* equal to $\overline{r_1}^{\mathrm{t}}, \dots, \overline{r_m}^{\mathrm{t}}$, and since entrywise conjugation carries a relation $\sum \lambda_k r_k = 0$ to $\sum \overline{\lambda_k}\,\overline{r_k} = 0$ and back, while transposition is an isomorphism of $\mathbf{F}^{1,n}$ onto $\mathbf{F}^{n,1}$,

$$\text{column rank of } A^* = \dim \operatorname{span}(\overline{r_1}, \dots, \overline{r_m}) = \dim \operatorname{span}(r_1, \dots, r_m),$$

which is the row rank of A. Hence the row and column ranks of A agree.

Problem 7A.9. Prove that the product of two self-adjoint operators on V is self-adjoint if and only if the two operators commute.

Solution. Suppose $S, T \in \mathcal{L}(V)$ are self-adjoint, so $S^* = S$ and $T^* = T$. By 7.5(d),

$$(ST)^* = T^* S^* = TS.$$

Hence $ST = (ST)^*$ if and only if $ST = TS$; that is, ST is self-adjoint exactly when S and T commute.

Problem 7A.10. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if

$$\langle Tv, v\rangle = \langle T^* v, v\rangle$$

for all $v \in V$.

Solution. By additivity of the inner product in its first slot,

$$\langle Tv, v\rangle - \langle T^* v, v\rangle = \big\langle (T - T^*)v,\, v\big\rangle$$

for every $v \in V$. Since $\mathbf{F} = \mathbf{C}$, result 7.13 applied to $T - T^*$ says that the right side vanishes for all $v \in V$ if and only if $T - T^* = 0$. Hence $\langle Tv, v\rangle = \langle T^* v, v\rangle$ for all $v \in V$ if and only if $T^* = T$.
Method (2): by 7.15, $\langle T^* v, v\rangle = \overline{\langle Tv, v\rangle}$, so the displayed condition says $\langle Tv, v\rangle \in \mathbf{R}$ for every v, which by 7.14 (valid as $\mathbf{F} = \mathbf{C}$) holds exactly when T is self-adjoint.

Problem 7A.11. Define an operator $S\colon \mathbf{F}^2 \to \mathbf{F}^2$ by $S(w, z) = (-z, w)$.
(a) Find a formula for S^*.
(b) Show that S is normal but not self-adjoint.
(c) Find all eigenvalues of S.
[If $\mathbf{F} = \mathbf{R}$, then S is the operator on $\mathbf{R}^2$ of counterclockwise rotation by 90°.]

Solution. (a) $S^*(w, z) = (z, -w)$.
Indeed, with $\mathbf{F}^2$ carrying its standard inner product (for which the standard basis is orthonormal), $\mathcal{M}(S)$ has columns $Se_1 = e_2$ and $Se_2 = -e_1$, so 7.9 gives

$$\mathcal{M}(S^*) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

the matrix of $(w, z) \mapsto (z, -w)$.
(b) S is normal because

$$SS^*(w, z) = S(z, -w) = (w, z) = S^*(-z, w) = S^* S(w, z),$$

so $SS^* = I = S^* S$; it is not self-adjoint because $S(1, 0) = (0, 1)$ while $S^*(1, 0) = (0, -1)$.
(c) $S(w, z) = \lambda(w, z)$ means $-z = \lambda w$ and $w = \lambda z$, whence $-z = \lambda^2 z$; an eigenvector must have $z \neq 0$ (else $w = \lambda z = 0$ too), so $\lambda^2 = -1$.
(i) If $\mathbf{F} = \mathbf{R}$, then S has no eigenvalues.
(ii) If $\mathbf{F} = \mathbf{C}$, then the eigenvalues are i and $-i$, realized by

$$S(i, 1) = i\,(i, 1), \qquad S(-i, 1) = -i\,(-i, 1).$$

Problem 7A.12. An operator $B \in \mathcal{L}(V)$ is called *skew* if

$$B^* = -B.$$

Suppose that $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exist commuting operators A and B such that A is self-adjoint, B is a skew operator, and $T = A + B$.

Solution. ($\Longrightarrow$) Take
$$A = \frac{T + T^*}{2}, \qquad B = \frac{T - T^*}{2},$$
so that $A + B = T$, and by 7.5(a), 7.5(b), 7.5(c) (the scalars $\pm\frac{1}{2}$ being real), $A^* = A$ and $B^* = -B$. They commute: expanding,
$$\begin{aligned} AB &= \frac{(T + T^*)(T - T^*)}{4} = \frac{T^2 - TT^* + T^*T - (T^*)^2}{4}, \\ BA &= \frac{(T - T^*)(T + T^*)}{4} = \frac{T^2 + TT^* - T^*T - (T^*)^2}{4}. \end{aligned}$$
so that $AB - BA = (T^*T - TT^*)/2 = 0$ by normality.
($\Longleftarrow$) Suppose $T = A + B$ with $A^* = A$, $B^* = -B$, and $AB = BA$. Then $T^* = A - B$ by 7.5(a) and 7.5(b), so
$$\begin{aligned} TT^* &= (A + B)(A - B) = A^2 - AB + BA - B^2 = A^2 - B^2, \\ T^*T &= (A - B)(A + B) = A^2 + AB - BA - B^2 = A^2 - B^2, \end{aligned}$$
the middle terms cancelling in both lines because $AB = BA$. Hence $TT^* = T^*T$, so T is normal.

Problem 7A.13. Suppose $\mathbf{F} = \mathbf{R}$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by $\mathcal{A}T = T^*$ for all $T \in \mathcal{L}(V)$.
(a) Find all eigenvalues of $\mathcal{A}$.
(b) Find the minimal polynomial of $\mathcal{A}$.

Solution. (a) The eigenvalues are 1 and -1 if $\dim V \geq 2$, and only 1 if $\dim V = 1$.
Here $\mathcal{A}$ is linear because $\mathbf{F} = \mathbf{R}$ (7.5(a) and 7.5(b), the conjugation in the latter being trivial), and 7.5(c) gives $\mathcal{A}^2 = I$. So if $\mathcal{A}T = \lambda T$ with $T \neq 0$, then
$$T = \mathcal{A}^2 T = \lambda^2 T,$$
forcing $\lambda \in \{1, -1\}$; the corresponding eigenvectors are the nonzero self-adjoint and the nonzero skew operators.
(i) $\dim V \geq 2$. Both occur: I_V is self-adjoint and nonzero, and for an orthonormal basis $e_1, \ldots, e_n$ of V the nonzero operator B with $Be_1 = e_2$, $Be_2 = -e_1$, $Be_k = 0$ for $k \geq 3$ has $\mathcal{M}(B)^* = -\mathcal{M}(B)$, hence $B^* = -B$ by 7.9.
(ii) $\dim V = 1$. Every $T \in \mathcal{L}(V)$ is λI with $\lambda \in \mathbf{R}$, hence self-adjoint by 7.5(b) and 7.5(e); so $\mathcal{A} = I$ and the only eigenvalue is 1.
(b) The minimal polynomial is $z^2 - 1$ if $\dim V \geq 2$, and $z - 1$ if $\dim V = 1$.
Since $\mathcal{A}^2 - I = 0$, by 5.29 the minimal polynomial divides $(z-1)(z+1)$. In case (i), 5.27 makes both 1 and -1 zeros of it, so it is $z^2 - 1$; in case (ii), $\mathcal{A} = I$ gives $z - 1$.

Problem 7A.14. Define an inner product on $\mathcal{P}_2(\mathbf{R})$ by $\langle p, q \rangle = \int_0^1 pq$. Define an operator $T \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ by
$$T(ax^2 + bx + c) = bx.$$
(a) Show that with this inner product, the operator T is not self-adjoint.
(b) The matrix of T with respect to the basis $1, x, x^2$ is
$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This matrix equals its conjugate transpose, even though T is not self-adjoint. Explain why this is not a contradiction.

Solution. (a) Take $p = x$ and $q = 1$, so $Tp = x$ and $Tq = 0$. Then

$$\langle Tp, q\rangle = \int_0^1 x\,dx = \frac{1}{2} \neq 0 = \langle p, Tq\rangle,$$

so $\langle Tp, q\rangle \neq \langle p, Tq\rangle$ and T is not self-adjoint (7.10 with 7.1).
(b) Because 7.9 requires an *orthonormal* basis, and $1, x, x^2$ is not orthonormal for this inner product:

$$\langle 1, x\rangle = \int_0^1 x\,dx = \frac{1}{2} \neq 0.$$

With respect to a nonorthonormal basis the matrix of T^* need not be the conjugate transpose of the matrix of T, so the symmetry of the displayed matrix carries no information about self-adjointness.

Problem 7A.15. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that
- (a) T is self-adjoint $\iff$ T^{-1} is self-adjoint;
- (b) T is normal $\iff$ T^{-1} is normal.

Solution. Both parts rest on 7.5(f): T^* is invertible with $(T^*)^{-1} = (T^{-1})^*$. In each part the converse follows by applying the forward implication to the invertible operator T^{-1}, since $(T^{-1})^{-1} = T$.
(a) If $T^* = T$, then

$$(T^{-1})^* = (T^*)^{-1} = T^{-1},$$

so T^{-1} is self-adjoint.
(b) If $TT^* = T^*T$, then taking inverses of these (invertible) operators and using $(AB)^{-1} = B^{-1}A^{-1}$ and 7.5(f) gives

$$(T^{-1})^*T^{-1} = (T^*)^{-1}T^{-1} = T^{-1}(T^*)^{-1} = T^{-1}(T^{-1})^*,$$

so T^{-1} is normal.

Problem 7A.16. Suppose $\mathbf{F} = \mathbf{R}$.
- (a) Show that the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.
- (b) What is the dimension of the subspace of $\mathcal{L}(V)$ in (a) [in terms of $\dim V$]?

Solution. Write $n = \dim V$ and $\mathcal{S} = \{T \in \mathcal{L}(V) : T^* = T\}$.
(a) $\mathcal{S}$ contains 0, and for $S, T \in \mathcal{S}$ and $\lambda \in \mathbf{R}$, results 7.5(a) and 7.5(b) give

$$(S + T)^* = S^* + T^* = S + T, \qquad (\lambda T)^* = \overline{\lambda}\,T^* = \lambda T,$$

the last equality because λ is real. Hence $\mathcal{S}$ is a subspace.
(b) $\dim \mathcal{S} = \dfrac{(\dim V)(\dim V + 1)}{2}$.
Fix an orthonormal basis of V (6.35) and take all matrices with respect to it. Then $\mathcal{M}$ is an isomorphism of $\mathcal{L}(V)$ onto $\mathbf{R}^{n,n}$ (3.71), and 7.9 gives $\mathcal{M}(T^*) = \mathcal{M}(T)^{\mathrm{t}}$ (conjugation being trivial over $\mathbf{R}$), so $\mathcal{M}$ restricts to an isomorphism of $\mathcal{S}$ onto the space of symmetric n-by-n real matrices. That space has basis

$$E_{j,j}\ (1 \le j \le n), \qquad E_{j,k} + E_{k,j}\ (1 \le j < k \le n)$$

(spanning and independence: read off entries; Check!), a list of length

$$n + \binom{n}{2} = \frac{n(n+1)}{2}.$$

Problem 7A.17. Suppose $\mathbf{F} = \mathbf{C}$. Show that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

Solution. The set is not closed under scalar multiplication: I is self-adjoint by 7.5(e), while by 7.5(b) and 7.5(e)

$$(iI)^* = \bar{i}\, I^* = -iI \neq iI,$$

the inequality because $I \neq 0$ (as $V \neq \{0\}$). So iI is not self-adjoint, and the set of self-adjoint operators is not a subspace of $\mathcal{L}(V)$.

Problem 7A.18. Suppose $\dim V \geq 2$. Show that the set of normal operators on V is not a subspace of $\mathcal{L}(V)$.

Solution. Fix an orthonormal basis $e_1, \ldots, e_n$ of V (6.35) and let $S, T \in \mathcal{L}(V)$ be the operators killing e_j for $j \geq 3$ whose matrices with respect to this basis have upper-left 2-by-2 blocks

$$\mathcal{M}(S) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathcal{M}(T) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

all other entries 0. Both are normal: since the basis is orthonormal, 7.9 and the injectivity of $\mathcal{M}$ (3.71) give $S^* = S$ (the matrix is real symmetric) and $T^* = -T$, and

$$TT^* = -T^2 = T^*T.$$

But $S + T$ is not normal: $(S + T)e_1 = 0$ while $\mathcal{M}\big((S + T)^*\big) = \mathcal{M}(S + T)^*$ gives $(S + T)^* e_1 = 2e_2$, so

$$\|(S + T)e_1\| = 0 \neq 2 = \|(S + T)^* e_1\|,$$

and 7.20 applies. Hence the set of normal operators is not closed under addition, so it is not a subspace of $\mathcal{L}(V)$.

Problem 7A.19. Suppose $T \in \mathcal{L}(V)$ and $\|T^*v\| \leq \|Tv\|$ for every $v \in V$. Prove that T is normal. [This exercise fails on infinite-dimensional inner product spaces, leading to what are called hyponormal operators, which have a well-developed theory.]

Solution. By 7.20 it suffices to prove $\|Tv\| = \|T^*v\|$ for every $v \in V$; the hypothesis becomes an equality because no slack is available over an orthonormal basis.
Let $e_1, \ldots, e_n$ be any orthonormal basis of V. Exercise 7A.5 (with $W = V$ and both bases equal to $e_1, \ldots, e_n$) gives

$$\sum_{k=1}^{n} \|Te_k\|^2 = \sum_{k=1}^{n} \|T^*e_k\|^2.$$

Each term $\|Te_k\|^2 - \|T^*e_k\|^2$ is nonnegative by hypothesis and they sum to 0, so

$$\|Te_k\| = \|T^*e_k\| \qquad \text{for each } k.$$

Now let $v \in V$ be nonzero (the case $v = 0$ being trivial) and put $e = v/\|v\|$, a unit vector, which by 6.36 extends to an orthonormal basis of V. The previous display applied to that basis gives $\|Te\| = \|T^*e\|$, and by homogeneity of the norm

$$\|Tv\| = \|v\|\, \|Te\| = \|v\|\, \|T^*e\| = \|T^*v\|.$$

Hence T is normal by 7.20.

Problem 7A.20. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the following are equivalent.
- (a) P is self-adjoint.
- (b) P is normal.
- (c) There is a subspace U of V such that $P = P_U$.

Solution. We prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).
(a) $\Rightarrow$ (b). If $P^* = P$, then $PP^* = P^2 = P^*P$, so P is normal.
(b) $\Rightarrow$ (c). Take $U = \operatorname{range} P$. By 7.21(a) and 7.6(a),

$$\operatorname{null} P = \operatorname{null} P^* = (\operatorname{range} P)^\perp = U^\perp,$$

while $P^2 = P$ makes P the identity on U: each $u \in U$ is $u = Px$, so $Pu = P^2x = Px = u$. Writing $v \in V$ as $v = u + w$ with $u \in U$, $w \in U^\perp$ (6.49), which is the decomposition defining $P_U v = u$ in 6.55, we get

$$Pv = Pu + Pw = u + 0 = P_U v.$$

(c) $\Rightarrow$ (a). With $v_1 = u_1 + w_1$ and $v_2 = u_2 + w_2$ decomposed as above,

$$\langle P_U v_1, v_2 \rangle = \langle u_1, u_2 \rangle = \langle v_1, P_U v_2 \rangle,$$

since $\langle u_1, w_2 \rangle = 0 = \langle w_1, u_2 \rangle$. Comparing with $\langle P_U v_1, v_2 \rangle = \langle v_1, P_U^* v_2 \rangle$ gives $\langle v_1, (P_U^* - P_U)v_2 \rangle = 0$ for all v_1, v_2; taking $v_1 = (P_U^* - P_U)v_2$ yields $P_U^* = P_U$.

Problem 7A.21. Suppose $D\colon \mathcal{P}_8(\mathbf{R}) \to \mathcal{P}_8(\mathbf{R})$ is the differentiation operator defined by $Dp = p'$. Prove that there does not exist an inner product on $\mathcal{P}_8(\mathbf{R})$ that makes D a normal operator.

Solution. The constant polynomial 1 lies in both $\operatorname{null} D$ and $\operatorname{range} D$, and these two subspaces are defined without reference to any inner product.
Indeed $D1 = 0$ and $Dx = 1$, so

$$1 \in (\operatorname{null} D) \cap (\operatorname{range} D).$$

If some inner product on $\mathcal{P}_8(\mathbf{R})$ (a nonzero finite-dimensional real vector space, so the standing hypotheses of this chapter hold) made D normal, then 7.21(c) would give

$$\mathcal{P}_8(\mathbf{R}) = \operatorname{null} D \oplus \operatorname{range} D,$$

forcing $(\operatorname{null} D) \cap (\operatorname{range} D) = \{0\}$ by 1.46, contradicting $1 \neq 0$.

Problem 7A.22. Give an example of an operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that T is normal but not self-adjoint.

Solution. Take $T(x, y, z) = (-y, x, z)$, rotation by 90° in the x, y-plane.
Its matrix with respect to the standard basis e_1, e_2, e_3, which is orthonormal, is

$$\mathcal{M}(T) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

By 7.9 the matrix of T^* in this orthonormal basis is the conjugate transpose, which here is the transpose, so $T^*(x, y, z) = (y, -x, z)$. Hence

$$TT^*(x, y, z) = T(y, -x, z) = (x, y, z) = T^*(-y, x, z) = T^*T(x, y, z),$$

so $TT^* = I = T^*T$ and T is normal (7.18). But $Te_1 = e_2$ while $T^*e_1 = -e_2$, so $T \neq T^*$.

Problem 7A.23. Suppose T is a normal operator on V. Suppose also that $v, w \in V$ satisfy the equations
$$\|v\| = \|w\| = 2, \quad Tv = 3v, \quad Tw = 4w.$$
Show that $\|T(v+w)\| = 10$.

Solution. Since v, w are nonzero (their norms are 2), they are eigenvectors of the normal operator T for the distinct eigenvalues 3 and 4, so $\langle v, w\rangle = 0$ by 7.22. Hence $3v$ and $4w$ are orthogonal, and the Pythagorean theorem (6.12) gives
$$\|T(v+w)\|^2 = \|3v+4w\|^2 = 9\|v\|^2 + 16\|w\|^2 = 36 + 64 = 100,$$
so $\|T(v+w)\| = 10$.

Problem 7A.24. Suppose $T \in \mathcal{L}(V)$ and
$$a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1} + z^m$$
is the minimal polynomial of T. Prove that the minimal polynomial of T^* is
$$\overline{a_0} + \overline{a_1} z + \overline{a_2} z^2 + \cdots + \overline{a_{m-1}} z^{m-1} + z^m.$$
[This exercise shows that the minimal polynomial of T^* equals the minimal polynomial of T if $\mathbf{F} = \mathbf{R}$.]

Solution. Write $\overline{q}$ for the polynomial obtained by conjugating each coefficient of q; thus $\overline{\overline{q}} = q$, $\deg \overline{q} = \deg q$, and $\overline{q}$ is monic exactly when q is.
The key identity is that $\big(q(S)\big)^* = \overline{q}(S^*)$ for all $S \in \mathcal{L}(V)$ and $q \in \mathcal{P}(\mathbf{F})$. Indeed $(S^k)^* = (S^*)^k$ by induction from $I^* = I$ (7.5(e)) and 7.5(d), so with $q(z) = \sum_k c_k z^k$, parts (a) and (b) of 7.5 give
$$\big(q(S)\big)^* = \sum_k \overline{c_k}\,(S^k)^* = \sum_k \overline{c_k}(S^*)^k = \overline{q}(S^*).$$
Let p be the minimal polynomial of T (monic of degree m) and q that of T^*; by 5.22 each is the monic polynomial of smallest degree annihilating its operator. Since $0^* = 0$, the identity with $S = T$ gives $\overline{p}(T^*) = \big(p(T)\big)^* = 0$, whence $\deg q \le m$; and with $S = T^*$, using $(T^*)^* = T$ from 7.5(c),
$$\overline{q}(T) = \big(q(T^*)\big)^* = 0,$$
whence $m \le \deg \overline{q} = \deg q$. So $\overline{q}$ is monic of degree m with $\overline{q}(T) = 0$, and the uniqueness in 5.22 forces $\overline{q} = p$. Conjugating coefficients,
$$q(z) = \overline{a_0} + \overline{a_1} z + \cdots + \overline{a_{m-1}} z^{m-1} + z^m.$$

Problem 7A.25. Suppose $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if T^* is diagonalizable.

Solution. Diagonalizability is detected by the minimal polynomial (5.62), and coefficient-conjugation preserves the shape that 5.62 requires.
Write $\overline{q}$ for the coefficientwise conjugate, as in Exercise 7A.24. Conjugation is multiplicative, $\overline{qr} = \overline{q}\,\overline{r}$, since the coefficient $\sum_{j+k=n} b_j c_k$ of z^n in qr conjugates to $\sum_{j+k=n} \overline{b_j}\,\overline{c_k}$.
Suppose T is diagonalizable. By 5.62 its minimal polynomial is $p(z) = (z - \lambda_1)\cdots(z - \lambda_m)$ with $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$ distinct, so
$$\overline{p}(z) = (z - \overline{\lambda_1})\cdots(z - \overline{\lambda_m}),$$

whose roots $\overline{\lambda_k}$ again lie in $\mathbf{F}$ (automatic for $\mathbf{F} = \mathbf{C}$; for $\mathbf{F} = \mathbf{R}$ they equal λ_k) and are still distinct, conjugation being injective. By Exercise 7A.24 the minimal polynomial of T^* is $\overline{p}$, so 5.62 makes T^* diagonalizable.
Conversely, if T^* is diagonalizable, apply this to T^* and use $(T^*)^* = T$ (7.5(c)).

Problem 7A.26. Fix $u, x \in V$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u\rangle x$ for every $v \in V$.
(a) Prove that if V is a real vector space, then T is self-adjoint if and only if the list u, x is linearly dependent.
(b) Prove that T is normal if and only if the list u, x is linearly dependent.

Solution. The adjoint is $T^*w = \langle w, x\rangle u$, because for all $v, w \in V$

$$\langle Tv, w\rangle = \langle v, u\rangle\langle x, w\rangle = \langle v, \langle w, x\rangle u\rangle.$$

Also u, x is linearly dependent if and only if $u = 0$ or $x = cu$ for some $c \in \mathbf{F}$. (Check!)
(a) If $u = 0$ then $T = 0 = T^*$. If $u \neq 0$ and $x = cu$ with $c \in \mathbf{R}$, then for every $v \in V$

$$Tv = c\langle v, u\rangle u = \langle v, cu\rangle u = T^*v$$

(here c real is what lets c pass out of the second slot), so T is self-adjoint. Conversely, if $T = T^*$ and $u \neq 0$, then evaluating at u gives $\|u\|^2 x = Tu = T^*u = \langle u, x\rangle u$, so $x = \big(\langle u, x\rangle/\|u\|^2\big)u$ and the list u, x is dependent (and if $u = 0$ it is dependent anyway).
(b) If $u = 0$ then $T = 0 = T^*$ is normal. If $u \neq 0$ and $x = cu$ with $c \in \mathbf{F}$, then for every $v \in V$, using $\langle x, x\rangle = |c|^2\|u\|^2$,

$$\begin{aligned} T^*Tv &= \langle v, u\rangle\langle x, x\rangle u = |c|^2\|u\|^2\langle v, u\rangle u,\\ TT^*v &= \langle v, x\rangle\langle u, u\rangle x = \|u\|^2\overline{c}c\langle v, u\rangle u, \end{aligned}$$

so $T^*T = TT^*$. Conversely, suppose T is normal and $u \neq 0$. Then $\|Tu\| = \|T^*u\|$ by 7.20, that is,

$$\|u\|^2\|x\| = \big\|\langle u, u\rangle x\big\| = \big\|\langle u, x\rangle u\big\| = |\langle u, x\rangle|\,\|u\|,$$

so $|\langle u, x\rangle| = \|u\|\,\|x\|$. This is equality in Cauchy-Schwarz, so by 6.14 one of u, x is a scalar multiple of the other and the list u, x is dependent.

Problem 7A.27. Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\operatorname{null} T^k = \operatorname{null} T \quad \text{and} \quad \operatorname{range} T^k = \operatorname{range} T$$

for every positive integer k.

Solution. Normality gives $V = \operatorname{null} T \oplus \operatorname{range} T$ by 7.21(c), hence

$$(\operatorname{null} T) \cap (\operatorname{range} T) = \{0\}$$

by 1.46; both conclusions follow from this.
(i) Null spaces. If $T^2v = 0$, then Tv lies in both $\operatorname{null} T$ and $\operatorname{range} T$, so $Tv = 0$; with the obvious reverse inclusion this gives $\operatorname{null} T^2 = \operatorname{null} T$. Inductively, if $\operatorname{null} T^k = \operatorname{null} T$ and $T^{k+1}v = 0$, then $Tv \in \operatorname{null} T^k = \operatorname{null} T$, so $v \in \operatorname{null} T^2 = \operatorname{null} T$. Hence $\operatorname{null} T^k = \operatorname{null} T$ for every $k \geq 1$.
(ii) Ranges. Since $T^kv = T(T^{k-1}v)$ we have $\operatorname{range} T^k \subseteq \operatorname{range} T$, while 3.21 applied to T^k and to T, together with (i), gives

$$\dim \operatorname{range} T^k = \dim V - \dim \operatorname{null} T = \dim \operatorname{range} T.$$

A subspace of the same dimension as the space containing it equals it (2.39), so $\operatorname{range} T^k = \operatorname{range} T$.

Problem 7A.28. Suppose $T \in \mathcal{L}(V)$ is normal. Prove that if $\lambda \in \mathbf{F}$, then the minimal polynomial of T is not a polynomial multiple of $(x-\lambda)^2$.

Solution. Suppose the minimal polynomial p of T satisfies $p(z) = (z-\lambda)^2 q(z)$ for some $q \in \mathcal{P}(\mathbf{F})$; we contradict the minimality of $\deg p$ in 5.22.
Comparing leading coefficients shows q is monic, so $r(z) = (z-\lambda)q(z)$ is monic with $\deg r = \deg p - 1$. Put $S = T - \lambda I$, normal by 7.21(d), so Exercise 7A.27 with $k = 2$ gives $\operatorname{null} S^2 = \operatorname{null} S$. For every $v \in V$, the multiplicativity 5.17(a) gives $p(T) = S^2 q(T)$ and $r(T) = S\,q(T)$, so

$$S^2\bigl(q(T)v\bigr) = p(T)v = 0,$$

putting $q(T)v \in \operatorname{null} S^2 = \operatorname{null} S$ and hence $r(T)v = S\bigl(q(T)v\bigr) = 0$. Thus r is monic with $r(T) = 0$ and $\deg r < \deg p$, the desired contradiction.

Problem 7A.29. Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \dots, e_n$ of V such that $\|Te_k\| = \|T^* e_k\|$ for each $k = 1, \dots, n$, then T is normal.

Solution. False: take $T(x,y) = (x+y, -x)$ on $\mathbf{F}^2$ with its standard inner product and standard orthonormal basis e_1, e_2.
For all $(x,y), (u,v) \in \mathbf{F}^2$,

$$\langle T(x,y), (u,v)\rangle = (x+y)\overline{u} - x\overline{v} = \bigl\langle (x,y), (u-v, u)\bigr\rangle,$$

so $T^*(u,v) = (u-v, u)$ by 7.1. Hence

$$Te_1 = (1,-1), \quad T^*e_1 = (1,1), \quad Te_2 = (1,0), \quad T^*e_2 = (-1,0),$$

giving $\|Te_k\| = \|T^*e_k\|$ for $k = 1, 2$. Yet

$$T^*Te_1 = T^*(1,-1) = (2,1) \neq (2,-1) = T(1,1) = TT^*e_1,$$

so $T^*T \neq TT^*$.

Problem 7A.30. Suppose that $T \in \mathcal{L}(\mathbf{F}^3)$ is normal and $T(1,1,1) = (2,2,2)$. Suppose $(z_1, z_2, z_3) \in \operatorname{null} T$. Prove that $z_1 + z_2 + z_3 = 0$.

Solution. With $u = (1,1,1)$ and the standard inner product, the assertion is exactly $\langle z, u\rangle = 0$.
The hypothesis says $Tu = 2u$, so 7.21(e) applied to the normal operator T gives $T^*u = \overline{2}\,u = 2u$. Hence for $z \in \operatorname{null} T$, the defining property of the adjoint (7.1) gives

$$0 = \langle Tz, u\rangle = \langle z, T^*u\rangle = \langle z, 2u\rangle = 2\langle z, u\rangle,$$

so $0 = \langle z, u\rangle = z_1 + z_2 + z_3$.

Problem 7A.31. Fix a positive integer n. In the inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product $\langle f, g\rangle = \int_{-\pi}^{\pi} fg$, let

$$V = \operatorname{span}(1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx).$$

(a) Define $D \in \mathcal{L}(V)$ by $Df = f'$. Show that $D^* = -D$. Conclude that D is normal but not self-adjoint.
(b) Define $T \in \mathcal{L}(V)$ by $Tf = f''$. Show that T is self-adjoint.

Solution. (a) Integration by parts gives $D^* = -D$, the boundary term vanishing because every $f \in V$ satisfies $f(-\pi) = f(\pi)$.
That periodicity holds for each of the $2n+1$ spanning functions (cosine is even, and $\sin(\pm k\pi) = 0$), hence for every linear combination. Also V is a finite-dimensional space invariant under differentiation, since $(\cos kx)' = -k \sin kx$ and $(\sin kx)' = k \cos kx$, so D is an operator on V and D^* exists. Elements of V are infinitely differentiable, so for $f, g \in V$,

$$\begin{aligned}\langle Df, g\rangle &= \int_{-\pi}^{\pi} f'g = \big[fg\big]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} fg' \\ &= f(\pi)g(\pi) - f(-\pi)g(-\pi) - \langle f, Dg\rangle = \langle f, (-D)g\rangle,\end{aligned}$$

and the adjoint is the unique operator with this property (7.1), so $D^* = -D$. Consequently

$$DD^* = -D^2 = D^*D,$$

so D is normal (7.18); and $D^* = D$ would give $2D = 0$, impossible since $D(\sin x) = \cos x \neq 0$.
(b) $T = D^2$ because V is closed under differentiation, so 7.5(d) and part (a) give

$$T^* = (DD)^* = D^*D^* = (-D)(-D) = D^2 = T.$$

Problem 7A.32. Suppose $T\colon V \to W$ is a linear map. Show that under the standard identification of V with V' (see 6.58) and the corresponding identification of W with W', the adjoint map $T^*\colon W \to V$ corresponds to the dual map $T'\colon W' \to V'$. More precisely, show that

$$T'(\varphi_w) = \varphi_{T^*w}$$

for all $w \in W$, where φ_w and φ_{T^*w} are defined as in 6.58.

Solution. Both $T'(\varphi_w)$ and φ_{T^*w} lie in V', so it suffices to evaluate them at an arbitrary $v \in V$. Fix $w \in W$; then

$$\begin{aligned}\big(T'(\varphi_w)\big)(v) &= (\varphi_w \circ T)(v) = \varphi_w(Tv) && \text{by 3.118} \\ &= \langle Tv, w\rangle && \text{by 6.58 in } W \\ &= \langle v, T^*w\rangle && \text{by 7.1} \\ &= \varphi_{T^*w}(v) && \text{by 6.58 in } V.\end{aligned}$$

As v was arbitrary, $T'(\varphi_w) = \varphi_{T^*w}$.

7.2 Exercises 7B

Problem 7B.1. Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.
This exercise strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.

Solution. If T is self-adjoint, then all its eigenvalues are real by 7.12 (normality is not needed here).
Conversely, suppose every eigenvalue of the normal operator T is real. Since $\mathbf{F} = \mathbf{C}$, the complex spectral theorem 7.31 gives an orthonormal basis $e_1, \dots, e_n$ of V with $Te_j = \lambda_j e_j$; each e_j is nonzero, so each λ_j is an eigenvalue and hence $\overline{\lambda_j} = \lambda_j$. Then 7.21(e) gives

$$T^*e_j = \overline{\lambda_j}\, e_j = \lambda_j e_j = Te_j$$

for each j, so T^* and T agree on a basis and $T^* = T$.

Problem 7B.2. Suppose $\mathbf{F} = \mathbf{C}$. Suppose $T \in \mathcal{L}(V)$ is normal and has only one eigenvalue. Prove that T is a scalar multiple of the identity operator.

Solution. $T = \lambda I$, where λ is the unique eigenvalue of T.
Since $\mathbf{F} = \mathbf{C}$ and T is normal, the complex spectral theorem 7.31 gives an orthonormal basis $e_1, \ldots, e_n$ of V with $Te_j = \lambda_j e_j$. Each e_j has norm 1, so each λ_j is an eigenvalue of T and hence $\lambda_j = \lambda$. Thus T and λI agree on a basis, so $T = \lambda I$.

Problem 7B.3. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that the set of eigenvalues of T is contained in $\{0, 1\}$ if and only if there is a subspace U of V such that $T = P_U$.

Solution. If $T = P_U$ and $Tv = \lambda v$ with $v \neq 0$, then $T^2 = T$ by 6.57(g) gives $\lambda^2 v = T^2 v = Tv = \lambda v$, so $\lambda^2 = \lambda$ and $\lambda \in \{0, 1\}$. (Normality is not needed here.)
Conversely, suppose every eigenvalue of T lies in $\{0, 1\}$. Since $\mathbf{F} = \mathbf{C}$ and T is normal, 7.31 gives an orthonormal basis $e_1, \ldots, e_n$ of V with $Te_j = \lambda_j e_j$, each λ_j an eigenvalue and so in $\{0, 1\}$; reorder so that $\lambda_1 = \cdots = \lambda_m = 1$ and $\lambda_{m+1} = \cdots = \lambda_n = 0$. Put $U = \operatorname{span}(e_1, \ldots, e_m)$, of which $e_1, \ldots, e_m$ is an orthonormal basis (take $U = \{0\}$ if $m = 0$). Expanding v in the orthonormal basis of V and applying T kills the terms beyond m:

$$Tv = \langle v, e_1 \rangle e_1 + \cdots + \langle v, e_m \rangle e_m = P_U v,$$

the last equality by 6.57(i). Hence $T = P_U$.

Problem 7B.4. Prove that a normal operator on a complex inner product space is skew (meaning it equals the negative of its adjoint) if and only if all its eigenvalues are purely imaginary (meaning that they have real part equal to 0).

Solution. If $T^* = -T$ and $Tv = \lambda v$ with $v \neq 0$, then

$$\lambda \|v\|^2 = \langle Tv, v \rangle = \langle v, T^* v \rangle = -\langle v, \lambda v \rangle = -\overline{\lambda} \|v\|^2,$$

so $\lambda = -\overline{\lambda}$ and λ is purely imaginary. (Normality is not needed here.)
Conversely, suppose every eigenvalue of the normal T is purely imaginary. Since $\mathbf{F} = \mathbf{C}$, the complex spectral theorem 7.31 gives an orthonormal basis $e_1, \ldots, e_n$ with $Te_j = \lambda_j e_j$; each λ_j is an eigenvalue, so $\lambda_j = ib_j$ with $b_j \in \mathbf{R}$ and $\overline{\lambda_j} = -\lambda_j$. Then 7.21(e) gives

$$T^* e_j = \overline{\lambda_j}\, e_j = -\lambda_j e_j = (-T) e_j$$

for each j, so $T^* = -T$.

Problem 7B.5. Prove or give a counterexample: If $T \in \mathcal{L}(\mathbf{C}^3)$ is a diagonalizable operator, then T is normal (with respect to the usual inner product).

Solution. False: take $T(z_1, z_2, z_3) = (z_1 + z_2,\ 2z_2,\ 3z_3)$ on $\mathbf{C}^3$.
With $u_1 = (1, 0, 0)$, $u_2 = (1, 1, 0)$, $u_3 = (0, 0, 1)$,

$$Tu_1 = u_1, \qquad Tu_2 = 2u_2, \qquad Tu_3 = 3u_3,$$

and u_1, u_2, u_3 is linearly independent (Check!), hence a basis of eigenvectors, so T is diagonalizable by 5.55. But u_1 and u_2 belong to the distinct eigenvalues 1 and 2 while $\langle u_1, u_2 \rangle = 1 \neq 0$, so by 7.22 the operator T is not normal.

Problem 7B.6. Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self-adjoint and $T^2 = T$.

Solution. Every eigenvalue of T lies in $\{0, 1\}$, and both conclusions follow from that.
Since $\mathbf{F} = \mathbf{C}$ and T is normal, the complex spectral theorem 7.31 gives an orthonormal basis $e_1, \dots, e_n$ of V with $Te_j = \lambda_j e_j$, whence $T^k e_j = \lambda_j^k e_j$ for every $k \geq 0$ by induction. Applying $T^9 = T^8$ to $e_j \neq 0$ gives $\lambda_j^8(\lambda_j - 1) = 0$, so $\lambda_j \in \{0, 1\}$.
(i) Each λ_j is real, so 7.21(e) gives $T^* e_j = \overline{\lambda_j} e_j = Te_j$ for each j; hence $T^* = T$.
(ii) Each λ_j satisfies $\lambda_j^2 = \lambda_j$, so

$$T^2 e_j = \lambda_j^2 e_j = \lambda_j e_j = Te_j$$

for each j; hence $T^2 = T$.

Problem 7B.7. Give an example of an operator T on a complex vector space such that $T^9 = T^8$ but $T^2 \neq T$.

Solution. Take $T \in \mathcal{L}(\mathbf{C}^2)$ with $T(w, z) = (z, 0)$.
Then $T^2(w, z) = T(z, 0) = (0, 0)$, so $T^2 = 0$ and hence $T^k = 0$ for all $k \geq 2$; in particular $T^9 = 0 = T^8$.
But $T^2 \neq T$, since $T(0, 1) = (1, 0) \neq (0, 0) = T^2(0, 1)$.

Problem 7B.8. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if every eigenvector of T is also an eigenvector of T^*.

Solution. If T is normal and $Tv = \lambda v$ with $v \neq 0$, then $T^* v = \overline{\lambda} v$ by 7.21(e), so v is an eigenvector of T^*.
Conversely, suppose every eigenvector of T is an eigenvector of T^*. Since $\mathbf{F} = \mathbf{C}$, Schur's theorem 6.38 gives an orthonormal basis $e_1, \dots, e_n$ of V in which $A = \mathcal{M}\big(T, (e_1, \dots, e_n)\big)$ is upper triangular; write $a_{j,k} = A_{j,k}$, so $a_{j,k} = 0$ for $j > k$. By 7.9 the matrix of T^* in this orthonormal basis is A^*, so

$$T^* e_k = \sum_{j=1}^{n} \overline{a_{k,j}}\, e_j \qquad \text{for each } k.$$

Induct on k to show that $a_{k,j} = 0$ for all $j \neq k$, assuming this for all rows $i < k$. In $Te_k = \sum_j a_{j,k} e_j$ the terms with $j > k$ vanish by upper triangularity and those with $j < k$ by the induction hypothesis applied to row j, so $Te_k = a_{k,k} e_k$ and the nonzero vector e_k is an eigenvector of T. By hypothesis $T^* e_k = \mu e_k$ for some $\mu \in \mathbf{C}$; comparing with the display and using that $e_1, \dots, e_n$ is a basis gives $a_{k,j} = 0$ for all $j \neq k$.
So A is diagonal, with $Te_k = a_{k,k} e_k$ and $T^* e_k = \overline{a_{k,k}}\, e_k$, whence

$$TT^* e_k = |a_{k,k}|^2 e_k = T^* T e_k$$

for every k, and $TT^* = T^* T$.

Problem 7B.9. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exists a polynomial $p \in \mathcal{P}(\mathbf{C})$ such that $T^* = p(T)$.

Solution. If $T^* = p(T)$, then T commutes with every power of T, hence with $p(T)$, so $TT^* = T\,p(T) = p(T)\,T = T^* T$ and T is normal.
Conversely, suppose T is normal. Since $\mathbf{F} = \mathbf{C}$, the complex spectral theorem 7.31 gives an orthonormal basis $e_1, \dots, e_n$ of V with $Te_j = \alpha_j e_j$, and 7.21(e) gives $T^* e_j = \overline{\alpha_j}\, e_j$. Let $\lambda_1, \dots, \lambda_m$ be the distinct

values among $\alpha_1, \ldots, \alpha_n$, and take the Lagrange interpolant

$$p(z) = \sum_{k=1}^{m} \overline{\lambda_k} \prod_{j \neq k} \frac{z - \lambda_j}{\lambda_k - \lambda_j},$$

with an empty product read as 1; the denominators are nonzero since the λ_k are distinct, and evaluating at $z = \lambda_i$ kills every term but the i-th, so $p(\lambda_i) = \overline{\lambda_i}$. Since $Te_j = \alpha_j e_j$ gives $q(T)e_j = q(\alpha_j)e_j$ for every $q \in \mathcal{P}(\mathbf{C})$, and α_j is one of the λ_k,

$$p(T)e_j = p(\alpha_j)e_j = \overline{\alpha_j}\, e_j = T^* e_j$$

for each j, so $p(T) = T^*$.

Problem 7B.10. Suppose V is a complex inner product space. Prove that every normal operator on V has a square root.
[An operator $S \in \mathcal{L}(V)$ is called a square root of $T \in \mathcal{L}(V)$ if $S^2 = T$. We will discuss more about square roots of operators in Sections 7C and 8C.]

Solution. Take S to be the operator sending e_j to $\mu_j e_j$, where $e_1, \ldots, e_n$ is an orthonormal basis of eigenvectors of the normal operator T and $\mu_j^2 = \lambda_j$.
Such a basis exists by the complex spectral theorem 7.31, with $Te_j = \lambda_j e_j$. Every complex number has a complex square root: writing $\lambda_j = r_j(\cos\theta_j + i \sin\theta_j)$ with $r_j \geq 0$,

$$\mu_j = \sqrt{r_j}\Big(\cos\tfrac{\theta_j}{2} + i \sin\tfrac{\theta_j}{2}\Big)$$

satisfies $\mu_j^2 = \lambda_j$ by the double-angle formulas. The linear map lemma 3.4 supplies $S \in \mathcal{L}(V)$ with $Se_j = \mu_j e_j$, and then

$$S^2 e_j = \mu_j^2 e_j = \lambda_j e_j = Te_j$$

for each j, so $S^2 = T$.

Problem 7B.11. Prove that every self-adjoint operator on V has a cube root.
[An operator $S \in \mathcal{L}(V)$ is called a cube root of $T \in \mathcal{L}(V)$ if $S^3 = T$.]

Solution. Take S to be the operator sending e_j to $\mu_j e_j$, where $e_1, \ldots, e_n$ is an orthonormal basis of eigenvectors of the self-adjoint T, with real eigenvalues λ_j, and $\mu_j^3 = \lambda_j$.
Such a basis exists in either field. (i) If $\mathbf{F} = \mathbf{R}$, the real spectral theorem 7.29 supplies it, and the eigenvalues are real because the scalar field is. (ii) If $\mathbf{F} = \mathbf{C}$, then T is normal ($TT^* = TT = T^*T$), so the complex spectral theorem 7.31 supplies it, and eigenvalues of a self-adjoint operator are real by 7.12.
Every real number has a real cube root, since $t \mapsto t^3$ maps $\mathbf{R}$ onto $\mathbf{R}$, so choose $\mu_j \in \mathbf{R}$ with $\mu_j^3 = \lambda_j$ and let 3.4 supply $S \in \mathcal{L}(V)$ with $Se_j = \mu_j e_j$. Then

$$S^3 e_j = \mu_j^3 e_j = \lambda_j e_j = Te_j$$

for each j, so $S^3 = T$.

Problem 7B.12. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is normal. Prove that if S is an operator on V that commutes with T, then S commutes with T^*.
[The result in this exercise is called Fuglede's theorem.]

Solution. Since T is normal and $\mathbf{F} = \mathbf{C}$, Exercise 7B.9 supplies $p \in \mathcal{P}(\mathbf{C})$ with $T^* = p(T)$.
From $ST = TS$ an induction gives $ST^k = T^k S$ for every $k \geq 0$, hence $S\,q(T) = q(T)\,S$ for every

$q \in \mathcal{P}(\mathbf{C})$. Taking $q = p$ gives $ST^* = T^*S$.
Method (2): Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. By 7.31 the space V has a basis of eigenvectors of T, so T is diagonalizable and

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$$

by 5.55. Each $E(\lambda_k, T)$ is invariant under S, since $T(Sv) = S(Tv) = \lambda_k(Sv)$, and 7.21(e) makes T^* act on $E(\lambda_k, T)$ as multiplication by $\overline{\lambda_k}$. So for $v \in E(\lambda_k, T)$,

$$ST^*v = \overline{\lambda_k}\, Sv = T^*Sv,$$

the second equality because Sv too lies in $E(\lambda_k, T)$; summing over the direct-sum decomposition of an arbitrary $v \in V$ gives $ST^* = T^*S$.

Problem 7B.13. Without using the complex spectral theorem, use the version of Schur's theorem that applies to two commuting operators (take $\mathcal{E} = \{T, T^*\}$ in Exercise 20 in Section 6B) to give a different proof that if $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal, then T has a diagonal matrix with respect to some orthonormal basis of V.

Solution. Take $\mathcal{E} = \{T, T^*\}$, whose elements commute pairwise precisely because T is normal. Exercise 20 in Section 6B then gives an orthonormal basis $e_1, \ldots, e_n$ of V with respect to which both

$$A = \mathcal{M}\big(T, (e_1, \ldots, e_n)\big) \quad \text{and} \quad B = \mathcal{M}\big(T^*, (e_1, \ldots, e_n)\big)$$

are upper triangular. Since the basis is orthonormal, $B = A^*$ by 7.9, that is, $B_{j,k} = \overline{A_{k,j}}$. Fix $j \neq k$.
(i) $j > k$: then $A_{j,k} = 0$, as A is upper triangular.
(ii) $j < k$: then $B_{k,j} = 0$, as B is upper triangular, and $B_{k,j} = \overline{A_{j,k}}$, so $A_{j,k} = 0$.
Hence A is diagonal.

Problem 7B.14. Suppose $\mathbf{F} = \mathbf{R}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Solution. Both directions are the real spectral theorem 7.29.
($\Rightarrow$) A self-adjoint operator is normal ($T^*T = T^2 = TT^*$), so eigenvectors corresponding to distinct eigenvalues are orthogonal by 7.22. Since T is self-adjoint, 7.29 supplies an orthonormal basis of V consisting of eigenvectors of T; thus T is diagonalizable and 5.55 gives

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T).$$

($\Leftarrow$) Each $E(\lambda_k, T)$ is a finite-dimensional inner product space, so it has an orthonormal basis $e_1^{(k)}, \ldots, e_{d_k}^{(k)}$ by 6.35. Concatenating these lists over $k = 1, \ldots, m$ produces a basis of V, since V is the direct sum of the $E(\lambda_k, T)$. That basis is orthonormal: within a block by construction, and across blocks $k \neq l$ because $e_i^{(k)}$ and $e_j^{(l)}$ are eigenvectors for the distinct eigenvalues λ_k and λ_l, hence orthogonal by hypothesis. Every vector in it satisfies $Te_i^{(k)} = \lambda_k e_i^{(k)}$, so V has an orthonormal basis of eigenvectors of T, and 7.29 gives $T^* = T$.

Problem 7B.15. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Solution. Both directions are the complex spectral theorem 7.31.
($\Rightarrow$) If T is normal, then eigenvectors corresponding to distinct eigenvalues are orthogonal by 7.22, and 7.31 gives an orthonormal basis of V consisting of eigenvectors of T; thus T is diagonalizable and 5.55 gives

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T).$$

($\Leftarrow$) Gram–Schmidt 6.32 applied to a basis of each $E(\lambda_j, T)$ yields an orthonormal basis $e_1^{(j)}, \ldots, e_{d_j}^{(j)}$ of $E(\lambda_j, T)$, where $d_j = \dim E(\lambda_j, T)$. Concatenating these lists over $j = 1, \ldots, m$ gives a basis of V, because V is the direct sum of the $E(\lambda_j, T)$. It is orthonormal: within a block by construction, and across blocks $i \neq j$ because $e_r^{(i)}$ and $e_s^{(j)}$ are eigenvectors for the distinct eigenvalues λ_i and λ_j, hence orthogonal by hypothesis. Every vector in it is an eigenvector of T, so 7.31 shows T is normal.

Problem 7B.16. Suppose $\mathbf{F} = \mathbf{C}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting normal operators for all $S, T \in \mathcal{E}$.
[This exercise extends the complex spectral theorem to the context of a collection of commuting normal operators.]

Solution. ($\Rightarrow$) Say every element of $\mathcal{E}$ has a diagonal matrix with respect to the orthonormal basis $e_1, \ldots, e_n$. Each $T \in \mathcal{E}$ is then normal by 7.31, and for $S, T \in \mathcal{E}$ the diagonal matrices $\mathcal{M}(S)$ and $\mathcal{M}(T)$ commute, so 3.43 gives

$$\begin{aligned}\mathcal{M}(ST) &= \mathcal{M}(S)\mathcal{M}(T) \\ &= \mathcal{M}(T)\mathcal{M}(S) = \mathcal{M}(TS);\end{aligned}$$

operators with the same matrix with respect to a fixed basis are equal, so $ST = TS$.
($\Leftarrow$) Induct on $n = \dim V$, assuming $\mathcal{E} \neq \emptyset$ (otherwise any orthonormal basis works). For $n = 0$ the empty basis works and for $n = 1$ any unit vector does, since every 1-by-1 matrix is diagonal. Let $n \geq 1$ and assume the claim in all smaller dimensions.
(i) Every $T \in \mathcal{E}$ is a scalar multiple of I: any orthonormal basis of V (6.35) works.
(ii) Some $T \in \mathcal{E}$ is not a scalar multiple of I. Since T is normal, 7.31 makes T diagonalizable and 5.55 gives $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$ for the distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ of T. Here $m \geq 2$, since $m = 1$ would give $T = \lambda_1 I$; as each eigenspace has dimension at least 1 and the dimensions sum to n, each $U_j = E(\lambda_j, T)$ has $\dim U_j < n$. Each U_j is invariant under every $S \in \mathcal{E}$, because for $v \in U_j$

$$T(Sv) = (TS)v = (ST)v = S(\lambda_j v) = \lambda_j(Sv).$$

So $\mathcal{E}_j = \{\, S|_{U_j} : S \in \mathcal{E} \,\} \subseteq \mathcal{L}(U_j)$ is defined; its elements are normal by 7B.20(d) (each S is normal and U_j is invariant under S) and commute pairwise, since $(S|_{U_j})(S'|_{U_j})u = (SS')u = (S'S)u = (S'|_{U_j})(S|_{U_j})u$ for $u \in U_j$. By induction U_j has an orthonormal basis $e_1^{(j)}, \ldots, e_{d_j}^{(j)}$ with respect to which every element of $\mathcal{E}_j$ is diagonal; that is, each $e_r^{(j)}$ is an eigenvector of every $S \in \mathcal{E}$. Concatenating over j gives a basis of V (as $V = U_1 \oplus \cdots \oplus U_m$), orthonormal because vectors from blocks $i \neq j$ are eigenvectors of the normal operator T for the distinct eigenvalues λ_i and λ_j, hence orthogonal by 7.22. Every vector of this orthonormal basis is an eigenvector of every $S \in \mathcal{E}$, so every element of $\mathcal{E}$ has a diagonal matrix with respect to it.

Problem 7B.17. Suppose $\mathbf{F} = \mathbf{R}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting self-adjoint operators for all $S, T \in \mathcal{E}$.
[This exercise extends the real spectral theorem to the context of a collection of commuting self-

adjoint operators.]

Solution. (⇒) Say every element of $\mathcal{E}$ has a diagonal matrix with respect to the orthonormal basis $e_1, \dots, e_n$. Each $T \in \mathcal{E}$ is then self-adjoint by 7.29, and diagonal matrices commute, so 3.43 gives

$$\begin{aligned}\mathcal{M}(ST) &= \mathcal{M}(S)\mathcal{M}(T)\\ &= \mathcal{M}(T)\mathcal{M}(S) = \mathcal{M}(TS)\end{aligned}$$

for $S, T \in \mathcal{E}$; hence $ST = TS$, operators with equal matrices being equal.
(⇐) Run the induction of 7B.16 on $n = \dim V$, with 7.29 in place of 7.31 and 7B.19(b) in place of 7B.20(d). For $\mathcal{E} = \emptyset$ or $n \leq 1$, or when every $T \in \mathcal{E}$ is a scalar multiple of I, any orthonormal basis of V (6.35) works. Otherwise fix $T \in \mathcal{E}$ that is not a scalar multiple of I: being self-adjoint, T is diagonalizable by 7.29, so 5.55 gives $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$ for the distinct eigenvalues of T, with $m \geq 2$ (else $T = \lambda_1 I$) and hence $\dim U_j < n$ for each $U_j = E(\lambda_j, T)$. Each U_j is invariant under every $S \in \mathcal{E}$, since for $v \in U_j$

$$T(Sv) = (TS)v = (ST)v = S(\lambda_j v) = \lambda_j(Sv),$$

and the restrictions $S|_{U_j}$ are self-adjoint by 7B.19(b) and commute pairwise, since $(S|_{U_j})(S'|_{U_j})u = (SS')u = (S'S)u = (S'|_{U_j})(S|_{U_j})u$ for $u \in U_j$. By induction each U_j has an orthonormal basis of common eigenvectors of all $S \in \mathcal{E}$; concatenating over j gives a basis of V, orthonormal because vectors from blocks $i \neq j$ are eigenvectors of T for the distinct eigenvalues λ_i and λ_j, hence orthogonal by 7.22 (T is normal, as $T^*T = T^2 = TT^*$). Every element of $\mathcal{E}$ therefore has a diagonal matrix with respect to this orthonormal basis.

Problem 7B.18. Give an example of a real inner product space V, an operator $T \in \mathcal{L}(V)$, and real numbers b, c with $b^2 < 4c$ such that

$$T^2 + bT + cI$$

is not invertible.
[This exercise shows that the hypothesis that T is self-adjoint cannot be deleted in 7.26, even for real vector spaces.]

Solution. Take $V = \mathbf{R}^2$ with the Euclidean inner product, $T(x, y) = (-y, x)$, and $b = 0$, $c = 1$, so that $b^2 = 0 < 4 = 4c$. Then $T^2(x, y) = T(-y, x) = -(x, y)$, so $T^2 = -I$ and

$$T^2 + bT + cI = -I + I = 0,$$

which is not invertible since $\mathbf{R}^2 \neq \{0\}$. This does not contradict 7.26, because T is not self-adjoint: $\langle T(1,0), (0,1)\rangle = 1$ while $\langle (1,0), T(0,1)\rangle = -1$.

Problem 7B.19. Suppose $T \in \mathcal{L}(V)$ is self-adjoint and U is a subspace of V that is invariant under T.
(a) Prove that $U^\perp$ is invariant under T.
(b) Prove that $T|_U \in \mathcal{L}(U)$ is self-adjoint.
(c) Prove that $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.

Solution. (a) For $v \in U^\perp$ and $u \in U$,

$$\langle Tv, u\rangle = \langle v, T^*u\rangle = \langle v, Tu\rangle = 0,$$

the last step because $Tu \in U$ while $v \in U^\perp$. Hence $Tv \in U^\perp$.

(b) Invariance gives $T|_U \in \mathcal{L}(U)$, and for all $u, w \in U$,

$$\begin{aligned}\langle (T|_U)u, w\rangle &= \langle Tu, w\rangle \\ &= \langle u, Tw\rangle = \langle u, (T|_U)w\rangle,\end{aligned}$$

the inner product of V restricting to that of U. Since the adjoint on U is the unique operator satisfying this identity (7.1), $(T|_U)^* = T|_U$.
(c) By (a) the subspace $U^\perp$ is invariant under T, so (b) applied to $U^\perp$ shows $T|_{U^\perp}$ is self-adjoint.

Problem 7B.20. Suppose $T \in \mathcal{L}(V)$ is normal and U is a subspace of V that is invariant under T.
(a) Prove that $U^\perp$ is invariant under T.
(b) Prove that U is invariant under T^*.
(c) Prove that $(T|_U)^* = (T^*)|_U$.
(d) Prove that $T|_U \in \mathcal{L}(U)$ and $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ are normal operators.
[This exercise can be used to give yet another proof of the complex spectral theorem (use induction on $\dim V$ and the result that T has an eigenvector).]

Solution. Everything follows once the matrix of T is shown to be block diagonal. Put $m = \dim U$, $n = \dim V$, extend an orthonormal basis $e_1, \ldots, e_m$ of U to an orthonormal basis $e_1, \ldots, e_n$ of V (6.36), and note that $e_{m+1}, \ldots, e_n$ is an orthonormal list in $U^\perp$ of length $n - m = \dim U^\perp$ (6.51), hence an orthonormal basis of $U^\perp$. Let $A = \mathcal{M}(T, (e_1, \ldots, e_n))$. Invariance of U gives $A_{j,k} = 0$ whenever $k \le m < j$, so

$$A = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$

with $B = \mathcal{M}(T|_U, (e_1, \ldots, e_m))$ of size m-by-m.
Now $C = 0$. Orthonormality makes $\|Te_k\|^2$ the sum of the squared absolute values of the entries in column k of A, while $\mathcal{M}(T^*) = A^*$ (7.9) makes $\|T^*e_k\|^2$ that sum over row k of A. Normality gives $\|Te_k\| = \|T^*e_k\|$ for every k (7.20), so summing over $k = 1, \ldots, m$ and using the vanishing block below B,

$$\sum_{j,k \le m} |B_{j,k}|^2 = \sum_{j,k \le m} |B_{k,j}|^2 + \sum_{k \le m < j} |A_{k,j}|^2.$$

The first terms on the two sides agree, so every entry of C is 0; thus A and A^* are block diagonal with blocks B, D and B^*, D^*.
(a) For $k > m$, column k of A vanishes in rows $1, \ldots, m$, so $Te_k \in \operatorname{span}(e_{m+1}, \ldots, e_n) = U^\perp$; linearity then gives $Tv \in U^\perp$ for all $v \in U^\perp$.
(b) For $k \le m$, column k of A^* vanishes in rows $m+1, \ldots, n$, so $T^*e_k \in \operatorname{span}(e_1, \ldots, e_m) = U$; linearity gives $T^*u \in U$ for all $u \in U$.
(c) By (b) the map $(T^*)|_U$ is an operator on U, and for $u, w \in U$,

$$\begin{aligned}\langle (T|_U)u, w\rangle &= \langle Tu, w\rangle \\ &= \langle u, T^*w\rangle = \langle u, (T^*)|_U w\rangle,\end{aligned}$$

so $(T|_U)^* = (T^*)|_U$ by the uniqueness of the adjoint on U (7.1).
(d) For $u \in U$ we have $T^*u \in U$ by (b), so (c) and $TT^* = T^*T$ give

$$(T|_U)(T|_U)^*u = T(T^*u) = (T^*T)u = T^*(Tu) = (T|_U)^*(T|_U)u,$$

whence $T|_U$ is normal. By (a), $U^\perp$ is another T-invariant subspace, so the same argument applied to $U^\perp$ shows $T|_{U^\perp}$ is normal.

Problem 7B.21. Suppose that T is a self-adjoint operator on a finite-dimensional inner product space and that 2 and 3 are the only eigenvalues of T. Prove that

$$T^2 - 5T + 6I = 0.$$

Solution. Since T is self-adjoint, the spectral theorem (7.29 if $\mathbf{F} = \mathbf{R}$; 7.31 if $\mathbf{F} = \mathbf{C}$, self-adjoint operators being normal) gives a basis $e_1, \ldots, e_n$ of V with $Te_k = \lambda_k e_k$. Each λ_k is an eigenvalue of T, so $\lambda_k \in \{2, 3\}$ by hypothesis, and hence

$$\begin{aligned}(T^2 - 5T + 6I)e_k &= (\lambda_k^2 - 5\lambda_k + 6)\, e_k \\ &= (\lambda_k - 2)(\lambda_k - 3)\, e_k = 0\end{aligned}$$

for every k. An operator vanishing on a basis is the zero operator, so $T^2 - 5T + 6I = 0$.
Method (2): T is diagonalizable, so its minimal polynomial is $(z - \mu_1) \cdots (z - \mu_M)$ with $\mu_1, \ldots, \mu_M$ distinct (5.62); the zeros of the minimal polynomial are exactly the eigenvalues of T (5.27(a)), so that polynomial is $(z - 2)(z - 3) = z^2 - 5z + 6$, and applying it to T gives $T^2 - 5T + 6I = 0$.

Problem 7B.22. Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^3)$ such that 2 and 3 are the only eigenvalues of T and $T^2 - 5T + 6I \neq 0$.

Solution. Take $T(z_1, z_2, z_3) = (2z_1 + z_2,\ 2z_2,\ 3z_3)$, whose matrix with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

This matrix is upper triangular, so by 5.41 the eigenvalues of T are exactly its diagonal entries 2 and 3. Since $z^2 - 5z + 6 = (z - 2)(z - 3)$, with $(T - 2I)(z_1, z_2, z_3) = (z_2, 0, z_3)$ and $(T - 3I)(z_1, z_2, z_3) = (-z_1 + z_2, -z_2, 0)$ we get

$$\begin{aligned}(T^2 - 5T + 6I)(0, 1, 0) &= (T - 2I)(T - 3I)(0, 1, 0) \\ &= (T - 2I)(1, -1, 0) = (-1, 0, 0) \neq 0.\end{aligned}$$

Problem 7B.23. Suppose $T \in \mathcal{L}(V)$ is self-adjoint, $\lambda \in \mathbf{F}$, and $\varepsilon > 0$. Suppose there exists $v \in V$ such that $\|v\| = 1$ and

$$\|Tv - \lambda v\| < \varepsilon.$$

Prove that T has an eigenvalue λ' such that $|\lambda - \lambda'| < \varepsilon$.
[This exercise shows that for a self-adjoint operator, a number that is close to satisfying an equation that would make it an eigenvalue is close to an eigenvalue.]

Solution. Take for λ' an eigenvalue of T closest to λ. Since T is self-adjoint, the spectral theorem (7.29 if $\mathbf{F} = \mathbf{R}$; 7.31 if $\mathbf{F} = \mathbf{C}$, self-adjoint operators being normal) gives an orthonormal basis $e_1, \ldots, e_n$ of V with $Te_k = \lambda_k e_k$, where $n \geq 1$ because $\|v\| = 1$ forces $V \neq \{0\}$. Pick k_0 minimizing $|\lambda - \lambda_k|$ and set

$$\lambda' = \lambda_{k_0}, \qquad d = |\lambda - \lambda'| = \min\{|\lambda - \lambda_k| : 1 \leq k \leq n\}.$$

Expanding $v = a_1 e_1 + \cdots + a_n e_n$ with $a_k = \langle v, e_k \rangle$ and $\sum_k |a_k|^2 = \|v\|^2 = 1$ (6.30) gives $Tv - \lambda v = \sum_k (\lambda_k - \lambda) a_k e_k$, a sum of pairwise orthogonal vectors, so the Pythagorean theorem 6.12 and $|\lambda_k - \lambda| \geq d$ yield

$$\begin{aligned}\|Tv - \lambda v\|^2 &= \sum_{k=1}^{n} |\lambda_k - \lambda|^2 |a_k|^2 \\ &\geq d^2 \sum_{k=1}^{n} |a_k|^2 = d^2.\end{aligned}$$

Hence $|\lambda - \lambda'| = d \leq \|Tv - \lambda v\| < \varepsilon$.

Problem 7B.24. Suppose U is a finite-dimensional vector space and $T \in \mathcal{L}(U)$.
(a) Suppose $\mathbf{F} = \mathbf{R}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis equals its transpose.
(b) Suppose $\mathbf{F} = \mathbf{C}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis commutes with its conjugate transpose.
[This exercise adds another equivalence to the list of conditions equivalent to diagonalizability in 5.55.]

Solution. The trick is that U carries no inner product in advance, so we may install one after seeing the basis: given a basis $u_1, \ldots, u_n$ of U, set

$$\Big\langle \sum_{j=1}^{n} a_j u_j,\ \sum_{j=1}^{n} b_j u_j \Big\rangle = \sum_{j=1}^{n} a_j \overline{b_j},$$

the pullback of the Euclidean inner product on $\mathbf{F}^n$ along the isomorphism $\sum_j a_j u_j \mapsto (a_1, \ldots, a_n)$, so the axioms 6.2 hold (Check!) and $u_1, \ldots, u_n$ is orthonormal for it. Write A^{t} and A^* for transpose and conjugate transpose; by 7.9, $\mathcal{M}(S^*) = (\mathcal{M}(S))^*$ with respect to an orthonormal basis.
(a) If T is diagonalizable, then 5.55 gives a basis of eigenvectors, with respect to which $\mathcal{M}(T)$ is diagonal and so equals its transpose. Conversely, if $A = \mathcal{M}(T, (u_1, \ldots, u_n))$ satisfies $A = A^{\mathrm{t}}$, install the inner product above; then $\mathbf{F} = \mathbf{R}$ and 7.9 give

$$\mathcal{M}(T^*) = A^* = A^{\mathrm{t}} = A = \mathcal{M}(T),$$

so $T^* = T$. Hence 7.29 supplies a basis of eigenvectors of T, and T is diagonalizable by 5.55.
(b) If T is diagonalizable, take the basis of eigenvectors from 5.55: then $D = \mathcal{M}(T)$ is diagonal with entries $\lambda_1, \ldots, \lambda_n$, and $DD^* = D^*D$ is diagonal with entries $|\lambda_1|^2, \ldots, |\lambda_n|^2$. Conversely, if $A = \mathcal{M}(T, (u_1, \ldots, u_n))$ satisfies $AA^* = A^*A$, install the inner product above; then 7.9 and 3.43 give

$$\mathcal{M}(TT^*) = AA^* = A^*A = \mathcal{M}(T^*T),$$

so T is normal. Hence 7.31 supplies a basis of eigenvectors of T, and T is diagonalizable by 5.55.

Problem 7B.25. Suppose that $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Show that if $k \in \{1, \ldots, n\}$, then the pseudoinverse $T^\dagger$ satisfies the equation

$$T^\dagger e_k = \begin{cases} \dfrac{1}{\lambda_k} e_k & \text{if } \lambda_k \neq 0, \\ 0 & \text{if } \lambda_k = 0. \end{cases}$$

Solution. By 6.68, $T^\dagger w = \big(T|_{(\operatorname{null} T)^\perp}\big)^{-1} P_{\operatorname{range} T}\, w$, where $T|_{(\operatorname{null} T)^\perp}$ is invertible from $(\operatorname{null} T)^\perp$ onto $\operatorname{range} T$ (6.67), so everything turns on identifying those two subspaces. Put

$$N = \operatorname{span}\{e_k : \lambda_k = 0\}, \qquad R = \operatorname{span}\{e_k : \lambda_k \neq 0\}.$$

Since $T(a_1 e_1 + \cdots + a_n e_n) = a_1 \lambda_1 e_1 + \cdots + a_n \lambda_n e_n$ and $e_1, \ldots, e_n$ is a basis, $Tv = 0$ exactly when $a_k = 0$ for every k with $\lambda_k \neq 0$; hence $\operatorname{null} T = N$, and $\operatorname{range} T = R$ because $\operatorname{range} T$ is spanned by $\lambda_1 e_1, \ldots, \lambda_n e_n$. Orthonormality gives $R \subseteq N^\perp$, and the two sublists together form the basis, so $\dim R = n - \dim N = \dim N^\perp$ by 6.51; equal dimensions force $R = N^\perp$ (2.39) and then $N = R^\perp$ (6.52). Thus

$$\operatorname{range} T = (\operatorname{null} T)^\perp, \qquad \operatorname{null} T = (\operatorname{range} T)^\perp.$$

(i) $\lambda_k = 0$: then $e_k \in N = (\operatorname{range} T)^\perp$, so $P_{\operatorname{range} T} e_k = 0$ and $T^\dagger e_k = 0$.

(ii) $\lambda_k \neq 0$: then $e_k \in R = \operatorname{range} T$, so $P_{\operatorname{range} T} e_k = e_k$; and $\frac{1}{\lambda_k} e_k \in R = (\operatorname{null} T)^\perp$ satisfies $T\big(\frac{1}{\lambda_k} e_k\big) = e_k$, so $T^\dagger e_k = \frac{1}{\lambda_k} e_k$.

7.3 Exercises 7C

Problem 7C.1. Suppose $T \in \mathcal{L}(V)$. Prove that if both T and $-T$ are positive operators, then $T = 0$.

Solution. For every $v \in V$, positivity of T and of $-T$ gives

$$\langle Tv, v\rangle \geq 0 \quad \text{and} \quad \langle Tv, v\rangle = -\langle (-T)v, v\rangle \leq 0,$$

so $\langle Tv, v\rangle = 0$. As T is positive, 7.43 then gives $Tv = 0$. Hence $T = 0$.

Problem 7C.2. Suppose $T \in \mathcal{L}(\mathbf{F}^4)$ is the operator whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

Show that T is an invertible positive operator.

Solution. Everything follows from the identity

$$\begin{aligned}\langle Tv, v\rangle = |v_1|^2 + |v_1 - v_2|^2 + |v_2 - v_3|^2 \\ + |v_3 - v_4|^2 + |v_4|^2.\end{aligned}$$

To get it, write $A = \mathcal{M}(T)$ for the displayed matrix. Reading off its rows,

$$Tv = (2v_1 - v_2,\ -v_1 + 2v_2 - v_3,\ -v_2 + 2v_3 - v_4,\ -v_3 + 2v_4),$$

so pairing with v and using $z + \overline{z} = 2 \operatorname{Re} z$,

$$\begin{aligned}\langle Tv, v\rangle = 2\sum_{k=1}^{4} |v_k|^2 \\ - 2\sum_{k=1}^{3} \operatorname{Re}\big(v_k \overline{v_{k+1}}\big),\end{aligned}$$

which is the right side of the first display, since $|v_k - v_{k+1}|^2 = |v_k|^2 + |v_{k+1}|^2 - 2\operatorname{Re}(v_k\overline{v_{k+1}})$ and each $|v_k|^2$ then occurs exactly twice. (Check!)
T is self-adjoint: the standard basis is orthonormal, and A is real and symmetric, so 7.9 gives $\mathcal{M}(T^*) = A^* = A = \mathcal{M}(T)$ and hence $T^* = T$. The identity exhibits $\langle Tv, v\rangle$ as a sum of squared absolute values, so $\langle Tv, v\rangle \geq 0$ and T is positive. Finally, $Tv = 0$ forces $\langle Tv, v\rangle = 0$, hence $v_1 = 0$, $v_1 = v_2$, $v_2 = v_3$, $v_3 = v_4$, so $v = 0$; thus T is injective and therefore invertible by 3.65.

Problem 7C.3. Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is the operator whose matrix (with respect to the standard basis) consists of all 1's. Show that T is a positive operator.

Solution. Writing $s = v_1 + \cdots + v_n$, every row of $A = \mathcal{M}(T)$ consists of 1s, so $Tv = (s, s, \ldots, s)$ and

$$\begin{aligned}\langle Tv, v\rangle &= \sum_{j=1}^{n} s\,\overline{v_j}\\ &= s\,\overline{s} = |s|^2 \geq 0.\end{aligned}$$

Also A is real and symmetric, so $A^* = A$; as the standard basis is orthonormal, 7.9 gives $\mathcal{M}(T^*) = A^* = A = \mathcal{M}(T)$ and hence $T^* = T$. A self-adjoint operator with $\langle Tv, v\rangle \geq 0$ for all v is positive.

Problem 7C.4. Suppose n is an integer with $n > 1$. Show that there exists an n-by-n matrix A such that all of the entries of A are positive numbers and $A = A^*$, but the operator on $\mathbf{F}^n$ whose matrix (with respect to the standard basis) equals A is not a positive operator.

Solution. Take $A_{1,2} = A_{2,1} = 2$ and $A_{j,k} = 1$ for all other j, k (legitimate since $n > 1$); for $n = 2$,

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Every entry is 1 or 2, hence positive, and A is real and symmetric, so $A = A^*$. Yet for $v = e_1 - e_2$ the coordinates of Tv are $1 - 2 = -1$, then $2 - 1 = 1$, then $1 - 1 = 0$, so

$$\langle Tv, v\rangle = (-1)\cdot\overline{1} + 1\cdot\overline{(-1)} = -2 < 0$$

and T is not a positive operator.

Problem 7C.5. Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that T is a positive operator if and only if for every orthonormal basis $e_1, \ldots, e_n$ of V, all entries on the diagonal of $\mathcal{M}(T, (e_1, \ldots, e_n))$ are nonnegative numbers.

Solution. The diagonal entries are the numbers $\langle Te_k, e_k\rangle$: expanding $Te_k = \sum_j \langle Te_k, e_j\rangle e_j$ (6.30) identifies column k of $A = \mathcal{M}(T, (e_1, \ldots, e_n))$, so

$$A_{j,k} = \langle Te_k, e_j\rangle, \qquad A_{k,k} = \langle Te_k, e_k\rangle.$$

These are real numbers, automatically when $\mathbf{F} = \mathbf{R}$ and by 7.14 when $\mathbf{F} = \mathbf{C}$, since T is self-adjoint.
($\Rightarrow$) If T is positive, then $A_{k,k} = \langle Te_k, e_k\rangle \geq 0$ for every orthonormal basis and every k, by the definition 7.34.
($\Leftarrow$) Since T is self-adjoint, the spectral theorem (7.29 if $\mathbf{F} = \mathbf{R}$, 7.31 if $\mathbf{F} = \mathbf{C}$) gives an orthonormal basis $f_1, \ldots, f_n$ with $Tf_k = \lambda_k f_k$; with respect to it the matrix of T is diagonal with entries $\lambda_k = \langle Tf_k, f_k\rangle$, which the hypothesis applied to this basis makes nonnegative. That is condition (c) of 7.38, so T is positive.

Problem 7C.6. Prove that the sum of two positive operators on V is a positive operator.

Solution. Let $S, T \in \mathcal{L}(V)$ be positive. Then $S^* = S$ and $T^* = T$, so $(S + T)^* = S^* + T^* = S + T$ by 7.5(a), and for every $v \in V$,

$$\langle (S + T)v, v\rangle = \langle Sv, v\rangle + \langle Tv, v\rangle \geq 0$$

by additivity of the inner product in its first slot. Hence $S + T$ is a positive operator.

Problem 7C.7. Suppose $S \in \mathcal{L}(V)$ is an invertible positive operator and $T \in \mathcal{L}(V)$ is a positive operator. Prove that $S + T$ is invertible.

Solution. It suffices to prove $S+T$ injective, since V is finite-dimensional (3.65). So let $(S+T)v=0$; pairing with v gives

$$0=\langle (S+T)v,v\rangle=\langle Sv,v\rangle+\langle Tv,v\rangle,$$

a sum of two nonnegative reals, so $\langle Sv,v\rangle=0$. As S is positive, 7.43 gives $Sv=0$, and S is invertible, hence injective, so $v=0$. Thus $S+T$ is injective and therefore invertible.

Problem 7C.8. Suppose $T\in\mathcal{L}(V)$. Prove that T is a positive operator if and only if the pseudoinverse $T^\dagger$ is a positive operator.

Solution. One direction is the spectral formula 7B.25; the other then follows from the involution $(T^\dagger)^\dagger=T$.

Suppose T is positive. Being self-adjoint, T has an orthonormal basis of eigenvectors $e_1,\dots,e_n$, say $Te_k=\lambda_k e_k$ (7.29 if $\mathbf{F}=\mathbf{R}$; 7.31 if $\mathbf{F}=\mathbf{C}$, self-adjoint operators being normal), and $\lambda_k\geq 0$ by 7.38(b). By 7B.25, $T^\dagger e_k=e_k/\lambda_k$ when $\lambda_k\neq 0$ and $T^\dagger e_k=0$ otherwise, so the matrix of $T^\dagger$ with respect to this orthonormal basis is diagonal with nonnegative entries; hence $T^\dagger$ is positive by 7.38(c).

For the converse it suffices to show $(T^\dagger)^\dagger=T$, since then the paragraph above applied to the positive operator $T^\dagger$ makes $(T^\dagger)^\dagger=T$ positive. Recall 6.68: $T^\dagger w=\big(T|_{(\operatorname{null}T)^\perp}\big)^{-1}P_{\operatorname{range}T}w$, with $T|_{(\operatorname{null}T)^\perp}$ invertible from $(\operatorname{null}T)^\perp$ onto $\operatorname{range}T$ (6.67). If $w\in(\operatorname{range}T)^\perp$ then $T^\dagger w=0$; conversely $T^\dagger w=0$ forces $P_{\operatorname{range}T}w=0$ by injectivity, and $w-P_{\operatorname{range}T}w\in(\operatorname{range}T)^\perp$ (6.55), so $w\in(\operatorname{range}T)^\perp$. Thus $\operatorname{null}T^\dagger=(\operatorname{range}T)^\perp$ and $(\operatorname{null}T^\dagger)^\perp=\operatorname{range}T$ by 6.52. Also $\operatorname{range}T^\dagger=(\operatorname{null}T)^\perp$: every value of $T^\dagger$ lies there, and $T^\dagger(Tv)=P_{(\operatorname{null}T)^\perp}v=v$ for $v\in(\operatorname{null}T)^\perp$ by 6.69(c). On $(\operatorname{null}T^\dagger)^\perp=\operatorname{range}T$ the map $T^\dagger$ equals $\big(T|_{(\operatorname{null}T)^\perp}\big)^{-1}$, whose inverse is $T|_{(\operatorname{null}T)^\perp}$, so for $v\in V$,

$$\begin{aligned}(T^\dagger)^\dagger v&=\big(T^\dagger|_{(\operatorname{null}T^\dagger)^\perp}\big)^{-1}P_{\operatorname{range}T^\dagger}v\\&=T\big(P_{(\operatorname{null}T)^\perp}v\big)=Tv,\end{aligned}$$

the last step because $v-P_{(\operatorname{null}T)^\perp}v\in\operatorname{null}T$.

Problem 7C.9. Suppose $T\in\mathcal{L}(V)$ is a positive operator and $S\in\mathcal{L}(W,V)$. Prove that S^*TS is a positive operator on W.

Solution. Since $S^*\in\mathcal{L}(V,W)$, the composition S^*TS is an operator on W. It is self-adjoint, because $T^*=T$ and 7.5(d) twice with 7.5(c) give

$$(S^*TS)^*=S^*T^*(S^*)^*=S^*TS.$$

And for $w\in W$, the adjoint identity $\langle S^*v,w\rangle=\langle v,Sw\rangle$ with $v=TSw$ gives

$$\langle (S^*TS)w,w\rangle=\langle TSw,\,Sw\rangle\geq 0,$$

since T is positive and $Sw\in V$. Hence S^*TS is positive by 7.34.

Problem 7C.10. Suppose T is a positive operator on V. Suppose $v,w\in V$ are such that

$$Tv=w\quad\text{and}\quad Tw=v.$$

Prove that $v=w$.

Solution. Linearity and the two hypotheses give $T(v-w)=Tv-Tw=w-v=-(v-w)$, so

$$\langle T(v-w),\,v-w\rangle=-\|v-w\|^2.$$

The left side is nonnegative because T is positive (7.34), so $\|v-w\|=0$ and hence $v=w$.

Problem 7C.11. Suppose T is a positive operator on V and U is a subspace of V invariant under T. Prove that $T|_U \in \mathcal{L}(U)$ is a positive operator on U.

Solution. Invariance gives $T|_U \in \mathcal{L}(U)$, with U carrying the inner product of V. Since T is positive it is self-adjoint, so for $u_1, u_2 \in U$,

$$\begin{aligned}\langle (T|_U)u_1,\, u_2\rangle &= \langle Tu_1,\, u_2\rangle \\ &= \langle u_1,\, Tu_2\rangle = \langle u_1,\, (T|_U)u_2\rangle,\end{aligned}$$

the last equality because $u_2 \in U$; as the adjoint on U is the unique operator satisfying this identity (7.1), $(T|_U)^* = T|_U$. Moreover $\langle (T|_U)u, u\rangle = \langle Tu, u\rangle \geq 0$ for every $u \in U$. Hence $T|_U$ is a positive operator on U by 7.34.

Problem 7C.12. Suppose $T \in \mathcal{L}(V)$ is a positive operator. Prove that T^k is a positive operator for every positive integer k.

Solution. Positivity of T gives $T^* = T$, so repeated use of 7.5(d) makes every power self-adjoint: $(T^j)^* = (T^*)^j = T^j$ for $j \geq 0$ (trivially for $j = 0$, as $T^0 = I$). Nonnegativity splits by the parity of k.
(i) $k = 2m$. For $v \in V$,

$$\begin{aligned}\langle T^{2m}v,\, v\rangle &= \langle T^m v,\, (T^m)^* v\rangle \\ &= \|T^m v\|^2 \geq 0.\end{aligned}$$

(ii) $k = 2m+1$. For $v \in V$,

$$\begin{aligned}\langle T^{2m+1}v,\, v\rangle &= \langle T(T^m v),\, (T^m)^* v\rangle \\ &= \langle T(T^m v),\, T^m v\rangle \geq 0,\end{aligned}$$

the inequality because T is positive. Hence T^k is a positive operator by 7.34.

Problem 7C.13. Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $\alpha \in \mathbf{R}$.
(a) Prove that $T - \alpha I$ is a positive operator if and only if α is less than or equal to every eigenvalue of T.
(b) Prove that $\alpha I - T$ is a positive operator if and only if α is greater than or equal to every eigenvalue of T.

Solution. Both parts are the equivalence (a) $\Leftrightarrow$ (b) of 7.38 applied to a shifted operator, so two preliminaries serve both.
First, $T - \alpha I$ and $\alpha I - T$ are self-adjoint: $(\alpha I)^* = \overline{\alpha}\, I = \alpha I$ since α is real (7.5(b), 7.5(e)), so 7.5(a) gives

$$(T - \alpha I)^* = T - \alpha I, \qquad (\alpha I - T)^* = \alpha I - T.$$

Second, for $v \neq 0$ we have $(T - \alpha I)v = \mu v \iff Tv = (\mu + \alpha)v$ and $(\alpha I - T)v = \mu v \iff Tv = (\alpha - \mu)v$, so the eigenvalues of $T - \alpha I$ are exactly the numbers $\lambda - \alpha$, and those of $\alpha I - T$ exactly the numbers $\alpha - \lambda$, as λ ranges over the eigenvalues of T (all real by 7.12).
(a) By 7.38, $T - \alpha I$ is positive if and only if every one of its eigenvalues $\lambda - \alpha$ is nonnegative, that is, if and only if $\alpha \leq \lambda$ for every eigenvalue λ of T.
(b) By 7.38, $\alpha I - T$ is positive if and only if every one of its eigenvalues $\alpha - \lambda$ is nonnegative, that is, if and only if $\alpha \geq \lambda$ for every eigenvalue λ of T.

Problem 7C.14. Suppose T is a positive operator on V and $v_1, \ldots, v_m \in V$. Prove that

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle Tv_k, v_j\rangle \geq 0.$$

Solution. The double sum equals $\langle Tv, v\rangle$ for $v = v_1 + \cdots + v_m$. Indeed, by linearity of T and additivity of the inner product in each slot (6.6(a), 6.6(d); all coefficients are 1, so no conjugates appear),

$$\begin{aligned}\sum_{j=1}^{m}\sum_{k=1}^{m}\langle Tv_k, v_j\rangle &= \sum_{j=1}^{m}\Big\langle T\Big(\sum_{k=1}^{m} v_k\Big),\ v_j\Big\rangle \\ &= \Big\langle Tv,\ \sum_{j=1}^{m} v_j\Big\rangle = \langle Tv, v\rangle.\end{aligned}$$

Since T is positive and $v \in V$, we have $\langle Tv, v\rangle \geq 0$ by 7.34, which is the asserted inequality.

Problem 7C.15. Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that there exist positive operators $A, B \in \mathcal{L}(V)$ such that

$$T = A - B \quad \text{and} \quad \sqrt{T^*T} = A + B \quad \text{and} \quad AB = BA = 0.$$

Solution. Take $A, B \in \mathcal{L}(V)$ defined on the spectral basis by the positive and negative parts of the eigenvalues. Precisely: T is self-adjoint, so the spectral theorem (7.29 if $\mathbf{F} = \mathbf{R}$, 7.31 if $\mathbf{F} = \mathbf{C}$, T being normal) gives an orthonormal basis $e_1, \dots, e_n$ of V with $Te_k = \lambda_k e_k$, each λ_k real by 7.12; put $\lambda^+ = \max\{\lambda, 0\}$, $\lambda^- = \max\{-\lambda, 0\}$, and let (3.4)

$$Ae_k = \lambda_k^+ e_k, \qquad Be_k = \lambda_k^- e_k \qquad (k = 1, \dots, n).$$

The matrices of A and B with respect to this orthonormal basis are diagonal with nonnegative entries, so both are positive by 7.38 ((c) implies (a)).
For each real λ we have $\lambda^+ - \lambda^- = \lambda$, $\lambda^+ + \lambda^- = |\lambda|$, and $\lambda^+\lambda^- = 0$ (one factor vanishes). Hence for each k,

$$\begin{aligned}(A - B)e_k &= \lambda_k e_k = Te_k, \\ ABe_k &= \lambda_k^-\lambda_k^+ e_k = 0 = \lambda_k^+\lambda_k^- e_k = BAe_k,\end{aligned}$$

so $T = A - B$ and $AB = BA = 0$.
Finally $C = A + B$ satisfies $Ce_k = |\lambda_k| e_k$, so C is positive by 7.38, and

$$C^2 e_k = \lambda_k^2 e_k = T^2 e_k = T^*Te_k$$

using $T^* = T$. Thus C is a positive square root of T^*T, so $A + B = \sqrt{T^*T}$ by uniqueness (7.39, 7.40).

Problem 7C.16. Suppose T is a positive operator on V. Prove that

$$\operatorname{null}\sqrt{T} = \operatorname{null} T \quad \text{and} \quad \operatorname{range}\sqrt{T} = \operatorname{range} T.$$

Solution. Write $R = \sqrt{T}$, so R is positive, $R^* = R$, and $R^2 = T$ (7.39, 7.40).
Null spaces: if $Rv = 0$ then $Tv = R(Rv) = 0$; conversely if $Tv = 0$ then

$$0 = \langle Tv, v\rangle = \langle R^2 v, v\rangle = \langle Rv, R^* v\rangle = \|Rv\|^2,$$

whence $Rv = 0$. Thus $\operatorname{null}\sqrt{T} = \operatorname{null} T$.
Ranges: both T and R are self-adjoint, so 7.6(b) gives

$$\operatorname{range} T = \operatorname{range} T^* = (\operatorname{null} T)^\perp = (\operatorname{null} R)^\perp = \operatorname{range} R^* = \operatorname{range}\sqrt{T}.$$

Problem 7C.17. Suppose that $T \in \mathcal{L}(V)$ is a positive operator. Prove that there exists a polynomial p with real coefficients such that $\sqrt{T} = p(T)$.

Solution. Take p to be the Lagrange interpolation polynomial sending each eigenvalue of T to its square root.
Being positive, T is self-adjoint, so the spectral theorem (7.29 or 7.31, according as $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$) gives an orthonormal basis $e_1, \dots, e_n$ of V with $Te_k = \lambda_k e_k$, each $\lambda_k \geq 0$ by 7.38 ((a) implies (b)). The operator R with $Re_k = \sqrt{\lambda_k}\, e_k$ (3.4) has diagonal matrix with nonnegative entries with respect to this orthonormal basis, hence is positive (7.38), and $R^2 e_k = \lambda_k e_k = Te_k$; so by uniqueness of positive square roots (7.39),

$$\sqrt{T}\, e_k = \sqrt{\lambda_k}\, e_k \qquad (k = 1, \dots, n).$$

Let $\mu_1, \dots, \mu_m$ be the distinct values among $\lambda_1, \dots, \lambda_n$ and set

$$p(x) = \sum_{r=1}^{m} \sqrt{\mu_r} \prod_{\substack{l=1 \\ l \neq r}}^{m} \frac{x - \mu_l}{\mu_r - \mu_l},$$

an empty product being 1. The denominators are nonzero (the μ_r are distinct) and every number in sight is real, so p has real coefficients; and $p(\mu_i) = \sqrt{\mu_i}$ for each i, since the summand $r = i$ has product 1 while every other summand contains the factor $\mu_i - \mu_i = 0$.
Since $Te_k = \lambda_k e_k$ forces $T^s e_k = \lambda_k^s e_k$ and hence $q(T)e_k = q(\lambda_k)e_k$ for every polynomial q,

$$p(T)e_k = p(\lambda_k)e_k = \sqrt{\lambda_k}\, e_k = \sqrt{T}\, e_k.$$

The two operators agree on a basis, so $\sqrt{T} = p(T)$.

Problem 7C.18. Suppose S and T are positive operators on V. Prove that ST is a positive operator if and only if S and T commute.

Solution. Both S and T are self-adjoint, being positive.
(i) If ST is positive, then ST is self-adjoint, so by 7.5(d)

$$ST = (ST)^* = T^* S^* = TS.$$

(ii) Suppose $ST = TS$. Then $(ST)^* = T^* S^* = TS = ST$, so ST is self-adjoint. Let $R = \sqrt{S}$, so $R^* = R$ and $R^2 = S$; by Exercise 7C.17, $R = p(S)$ for a polynomial p, and T commutes with every polynomial in S, so $RT = TR$. Hence for $v \in V$,

$$\begin{aligned}\langle STv, v\rangle &= \langle R(RT)v, v\rangle = \langle R(TR)v, v\rangle \\ &= \langle T(Rv), R^* v\rangle = \langle T(Rv), Rv\rangle \geq 0\end{aligned}$$

since T is positive. So ST is positive by 7.34.

Problem 7C.19. Show that the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots.

Solution. For $t \in [0, \pi)$ take $R_t \in \mathcal{L}(\mathbf{F}^2)$ with matrix, relative to the (orthonormal) standard basis,

$$\mathcal{M}(R_t) = \begin{pmatrix} \cos t & \sin t \\ \sin t & -\cos t \end{pmatrix}.$$

This matrix is real and symmetric, hence its own conjugate transpose, so $R_t^* = R_t$ by 7.9 (the basis is orthonormal). Squaring it gives

$$\begin{pmatrix} \cos^2 t + \sin^2 t & 0 \\ 0 & \sin^2 t + \cos^2 t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

so $R_t^2 = I$ by 3.43.

The R_t are pairwise distinct: $R_t e_1 = (\cos t, \sin t)$, and $\cos s = \cos t$ with $\sin s = \sin t$ forces $s = t$ for $s, t \in [0, \pi)$. As $[0, \pi)$ is infinite, I has infinitely many self-adjoint square roots.

Problem 7C.20. Suppose $T \in \mathcal{L}(V)$ and $e_1, \ldots, e_n$ is an orthonormal basis of V. Prove that T is a positive operator if and only if there exist $v_1, \ldots, v_n \in V$ such that

$$\langle Te_k, e_j \rangle = \langle v_k, v_j \rangle$$

for all $j, k = 1, \ldots, n$.
[The numbers $\{\langle Te_k, e_j \rangle\}_{j,k=1,\ldots,n}$ are the entries in the matrix of T with respect to the orthonormal basis $e_1, \ldots, e_n$.]

Solution. The witnesses are $v_k = \sqrt{T}\, e_k$. Throughout we use that operators A, B with $\langle Ae_k, e_j \rangle = \langle Be_k, e_j \rangle$ for all j, k are equal, since then $Ae_k = \sum_j \langle Ae_k, e_j \rangle e_j = Be_k$ by 6.30(a).
(i) If T is positive, put $R = \sqrt{T}$, so $R^* = R$ and $R^2 = T$ (7.39, 7.40), and $v_k = Re_k$. Then for all j, k,

$$\langle v_k, v_j \rangle = \langle Re_k, Re_j \rangle = \langle R^* Re_k, e_j \rangle = \langle Te_k, e_j \rangle.$$

(ii) Conversely, given such $v_1, \ldots, v_n$, let $R \in \mathcal{L}(V)$ satisfy $Re_k = v_k$ (3.4). Then for all j, k,

$$\langle R^* Re_k, e_j \rangle = \langle Re_k, Re_j \rangle = \langle v_k, v_j \rangle = \langle Te_k, e_j \rangle,$$

so $T = R^* R$, which is positive by 7.38 ((f) implies (a)).

Problem 7C.21. Suppose n is a positive integer. The n-by-n Hilbert matrix is the n-by-n matrix whose entry in row j, column k is $\frac{1}{j+k-1}$. Suppose $T \in \mathcal{L}(V)$ is an operator whose matrix with respect to some orthonormal basis of V is the n-by-n Hilbert matrix. Prove that T is a positive invertible operator.
Example: The 4-by-4 Hilbert matrix is

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

Solution. Everything follows from the integral representation $\frac{1}{j+k-1} = \int_0^1 x^{k-1} x^{j-1}\, dx$ (valid since $j + k - 2 \geq 0$).
Let $e_1, \ldots, e_n$ be an orthonormal basis of V with $\langle Te_k, e_j \rangle = \frac{1}{j+k-1}$, the matrix entry in row j, column k being $\langle Te_k, e_j \rangle$ by 6.30(a).
Self-adjoint: the Hilbert matrix is real and symmetric, hence its own conjugate transpose, so $T^* = T$ by 7.9.
Positive: write $v = a_1 e_1 + \cdots + a_n e_n$ and $q(x) = a_1 + a_2 x + \cdots + a_n x^{n-1}$. By linearity in the first slot and conjugate-linearity in the second, then linearity of the integral (and x real, so $\overline{q(x)} = \sum_j \overline{a_j} x^{j-1}$),

$$\begin{aligned} \langle Tv, v \rangle &= \sum_{k=1}^{n} \sum_{j=1}^{n} a_k \overline{a_j} \int_0^1 x^{k-1} x^{j-1}\, dx \\ &= \int_0^1 q(x)\, \overline{q(x)}\, dx = \int_0^1 |q(x)|^2\, dx \ \geq 0, \end{aligned}$$

so T is positive by 7.34.
Invertible: if $Tv = 0$ then $\int_0^1 |q|^2 = \langle Tv, v \rangle = 0$; since $|q|^2$ is continuous and nonnegative, it vanishes identically on $[0, 1]$ (were $|q(x_0)|^2 = c > 0$, continuity would give $|q|^2 > c/2$ on a subinterval of length $\delta > 0$ and hence $\int_0^1 |q|^2 \geq c\delta/2 > 0$). So q has infinitely many zeros, forcing $a_1 = \cdots = a_n = 0$ by 4.8, i.e. $v = 0$. Hence T is injective (3.15) and therefore invertible (3.65).

Problem 7C.22. Suppose $T \in \mathcal{L}(V)$ is a positive operator and $u \in V$ is such that $\|u\| = 1$ and $\|Tu\| \geq \|Tv\|$ for all $v \in V$ with $\|v\| = 1$. Show that u is an eigenvector of T corresponding to the largest eigenvalue of T.

Solution. Expanding u in a spectral basis shows that only the top eigenvalue's coordinates survive. Since T is positive it is self-adjoint with nonnegative eigenvalues (7.38), so the spectral theorem (7.29 if $\mathbf{F} = \mathbf{R}$, 7.31 if $\mathbf{F} = \mathbf{C}$) gives an orthonormal basis $e_1, \dots, e_n$ with $Te_k = \lambda_k e_k$, each $\lambda_k \geq 0$. Every eigenvalue of T occurs among $\lambda_1, \dots, \lambda_n$: if $Tw = \lambda w$ with $w = \sum_k a_k e_k \neq 0$, then $a_k(\lambda_k - \lambda) = 0$ for each k, and some $a_k \neq 0$. Relabel so that $\lambda_1 = \max_k \lambda_k$, the largest eigenvalue.
Write $u = a_1 e_1 + \cdots + a_n e_n$; then $\sum_k |a_k|^2 = 1$ and $\|Tu\|^2 = \sum_k \lambda_k^2 |a_k|^2$ by 6.30. Applying the hypothesis to the unit vector e_1 gives $\|Tu\| \geq \|Te_1\| = \lambda_1$, so

$$\lambda_1^2 \leq \|Tu\|^2 = \sum_{k=1}^{n} \lambda_k^2 |a_k|^2 \leq \sum_{k=1}^{n} \lambda_1^2 |a_k|^2 = \lambda_1^2,$$

using $0 \leq \lambda_k \leq \lambda_1$. Hence $\sum_k (\lambda_1^2 - \lambda_k^2)|a_k|^2 = 0$ with every term nonnegative, so $a_k = 0$ whenever $\lambda_k \neq \lambda_1$ (for nonnegative numbers, $\lambda_k^2 = \lambda_1^2$ iff $\lambda_k = \lambda_1$). Thus u is a combination of eigenvectors all with eigenvalue λ_1, giving $Tu = \lambda_1 u$ with $u \neq 0$.

Problem 7C.23. For $T \in \mathcal{L}(V)$ and $u, v \in V$, define $\langle u, v\rangle_T$ by $\langle u, v\rangle_T = \langle Tu, v\rangle$.
(a) Suppose $T \in \mathcal{L}(V)$. Prove that $\langle \cdot, \cdot\rangle_T$ is an inner product on V if and only if T is an invertible positive operator (with respect to the original inner product $\langle \cdot, \cdot\rangle$).
(b) Prove that every inner product on V is of the form $\langle \cdot, \cdot\rangle_T$ for some positive invertible operator $T \in \mathcal{L}(V)$.

Solution. Throughout, adjoints, positivity and self-adjointness refer to the original inner product $\langle \cdot, \cdot\rangle$.
(a) (i) Let T be an invertible positive operator; we check the conditions of 6.2. Positivity and additivity and homogeneity in the first slot are immediate from positivity and linearity of T (Check!). Definiteness: $\langle v, v\rangle_T = \langle Tv, v\rangle = 0$ forces $Tv = 0$ by 7.43, hence $v = 0$ since T is injective. Conjugate symmetry uses $T^* = T$:

$$\langle v, u\rangle_T = \langle Tv, u\rangle = \langle v, Tu\rangle = \overline{\langle Tu, v\rangle} = \overline{\langle u, v\rangle_T}.$$

(ii) Conversely, let $\langle \cdot, \cdot\rangle_T$ be an inner product. Conjugate symmetry of both inner products gives, for all u, v,

$$\langle u, T^*v\rangle = \langle Tu, v\rangle = \overline{\langle Tv, u\rangle} = \langle u, Tv\rangle,$$

the first equality by 7.1; taking $u = T^*v - Tv$ yields $T^* = T$. Positivity of $\langle \cdot, \cdot\rangle_T$ gives $\langle Tv, v\rangle \geq 0$, so T is positive (7.34). And $Tv = 0$ implies $\langle v, v\rangle_T = 0$, hence $v = 0$; an injective operator on V is invertible (3.65).
(b) Given an inner product $(\cdot, \cdot)$ on V, apply the Riesz representation theorem (6.42) for $\langle \cdot, \cdot\rangle$ to the linear functional $v \mapsto (v, u)$: there is a unique $Tu \in V$ with

$$(v, u) = \langle v, Tu\rangle \quad \text{for all } v \in V.$$

This T is linear: for all v,

$$\begin{aligned} \langle v, T(u_1 + u_2)\rangle &= (v, u_1) + (v, u_2) = \langle v, Tu_1 + Tu_2\rangle, \\ \langle v, T(\lambda u)\rangle &= \overline{\lambda}\,(v, u) = \langle v, \lambda Tu\rangle, \end{aligned}$$

by additivity and conjugate homogeneity in the second slot of both inner products; taking v to be the difference of the two arguments in each line gives $T \in \mathcal{L}(V)$. Then

$$\langle u, v\rangle_T = \langle Tu, v\rangle = \overline{\langle v, Tu\rangle} = \overline{(v, u)} = (u, v),$$

so $\langle \cdot, \cdot\rangle_T = (\cdot, \cdot)$ is an inner product and T is positive and invertible by (a).

Problem 7C.24. Suppose S and T are positive operators on V. Prove that

$$\operatorname{null}(S+T) = \operatorname{null} S \cap \operatorname{null} T.$$

Solution. The inclusion $\operatorname{null} S \cap \operatorname{null} T \subseteq \operatorname{null}(S+T)$ is immediate from $(S+T)v = Sv + Tv$. Conversely, if $v \in \operatorname{null}(S+T)$ then

$$0 = \langle Sv + Tv, v\rangle = \langle Sv, v\rangle + \langle Tv, v\rangle,$$

and both summands are nonnegative since S and T are positive, so both vanish. By 7.43 applied to S and to T, we get $Sv = Tv = 0$, i.e. $v \in \operatorname{null} S \cap \operatorname{null} T$.

Problem 7C.25. Let T be the second derivative operator in Exercise 31(b) in Section 7A. Show that $-T$ is a positive operator.

Solution. With $D \in \mathcal{L}(V)$ the differentiation operator of Exercise 7A.31(a), we have $-T = D^*D$, which is positive by 7.38 ((f) implies (a)).
Here $V = \operatorname{span}(1, \cos x, \dots, \cos nx, \sin x, \dots, \sin nx)$ inside the space of continuous real-valued functions on $[-\pi, \pi]$ with $\langle f, g\rangle = \int_{-\pi}^{\pi} fg$, and $Tf = f''$. Differentiation maps V into V, since

$$1' = 0, \qquad (\cos kx)' = -k \sin kx, \qquad (\sin kx)' = k \cos kx$$

all lie in V; so $D \in \mathcal{L}(V)$ and $T = D^2$.
By Exercise 7A.31(a), $D^* = -D$: every $f \in V$ is 2π-periodic and continuously differentiable, so the boundary term in

$$\langle Df, g\rangle = \Big[fg\Big]_{-\pi}^{\pi} - \int_{-\pi}^{\pi} fg' = -\langle f, Dg\rangle$$

vanishes. Hence $-T = -D^2 = (-D)D = D^*D$.

7.4 Exercises 7D

Problem 7D.1. Suppose $\dim V \geq 2$ and $S \in \mathcal{L}(V, W)$. Prove that S is an isometry if and only if Se_1, Se_2 is an orthonormal list in W for every orthonormal list e_1, e_2 of length two in V.

Solution. Write $n = \dim V \geq 2$.
(i) If S is an isometry then S preserves inner products by 7.49 ((a) implies (c)), so for any orthonormal e_1, e_2 and $j, k \in \{1, 2\}$,

$$\langle Se_j, Se_k\rangle = \langle e_j, e_k\rangle,$$

which says Se_1, Se_2 is orthonormal in W.
(ii) Conversely, take an orthonormal basis $e_1, \dots, e_n$ of V (6.35). For $j \neq k$ the pair e_j, e_k is an orthonormal list of length two, so by hypothesis $\|Se_j\| = \|Se_k\| = 1$ and $\langle Se_j, Se_k\rangle = 0$; since $n \geq 2$, every index j occurs in such a pair, so $Se_1, \dots, Se_n$ is orthonormal in W. Hence S is an isometry by 7.49 ((d) implies (a)).

Problem 7D.2. Suppose $T \in \mathcal{L}(V, W)$ and $T \neq 0$. Prove that T is a scalar multiple of an isometry if and only if T preserves orthogonality.
[The phrase "T preserves orthogonality" means that $\langle Tu, Tv\rangle = 0$ for all $u, v \in V$ such that $\langle u, v\rangle = 0$.]

Solution. (i) If $T = cS$ with S an isometry, then S preserves inner products by 7.49 ((a) implies (c)), so $\langle u, v\rangle = 0$ gives

$$\langle Tu, Tv\rangle = c\bar{c}\,\langle Su, Sv\rangle = |c|^2\langle u, v\rangle = 0.$$

(ii) Conversely, suppose T preserves orthogonality. The claim is that $\|Tu\|$ takes one value c on all unit vectors; then $c > 0$ (else $Tv = \|v\|\, T(v/\|v\|) = 0$ for all $v \neq 0$, contradicting $T \neq 0$) and $S = \frac{1}{c}T$ satisfies $\|Sv\| = \frac{\|v\|}{c}\|T(v/\|v\|)\| = \|v\|$ for $v \neq 0$ and trivially at 0, so $T = cS$ with S an isometry.
For the claim, first let u, v be orthonormal. Then $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2 = 0$, so $T(u+v) \perp T(u-v)$; also $\langle Tu, Tv\rangle = 0 = \langle Tv, Tu\rangle$. Expanding,

$$0 = \langle Tu + Tv,\, Tu - Tv\rangle = \|Tu\|^2 - \|Tv\|^2.$$

Now fix a unit vector u (possible as $T \neq 0$ forces $V \neq \{0\}$) and set $c = \|Tu\|$; let $\|v\| = 1$.
(a) If $v = \lambda u$ then $|\lambda| = 1$ and $\|Tv\| = |\lambda|\, \|Tu\| = c$.
(b) Otherwise $v - \langle v, u\rangle u \neq 0$; let e be its normalization, so u, e is orthonormal (Check!) and $\|Te\| = c$ by the previous paragraph, while $\langle Tu, Te\rangle = 0$. Rearranging gives $v \in \operatorname{span}(u, e)$, say $v = au + be$ with $|a|^2 + |b|^2 = \|v\|^2 = 1$ by the Pythagorean theorem (6.12). Applying T and using $Tu \perp Te$,

$$\|Tv\|^2 = |a|^2\|Tu\|^2 + |b|^2\|Te\|^2 = (|a|^2 + |b|^2)c^2 = c^2.$$

Problem 7D.3. (a) Show that the product of two unitary operators on V is a unitary operator.
(b) Show that the inverse of a unitary operator on V is a unitary operator.
[This exercise shows that the set of unitary operators on V is a group, where the group operation is the usual product of two operators.]

Solution. A unitary operator is an invertible isometry (7.51).
(a) If S, T are unitary then ST is invertible, with inverse $T^{-1}S^{-1}$, and for $v \in V$,

$$\|(ST)v\| = \|S(Tv)\| = \|Tv\| = \|v\|,$$

using the isometry of S at Tv and then that of T. So ST is unitary.
(b) If S is unitary then S^{-1} is invertible (inverse S), and for $v \in V$, setting $w = S^{-1}v$,

$$\|S^{-1}v\| = \|w\| = \|Sw\| = \|v\|,$$

so S^{-1} is a unitary operator.

Problem 7D.4. Suppose $\mathbf{F} = \mathbf{C}$ and $A, B \in \mathcal{L}(V)$ are self-adjoint. Show that $A + iB$ is unitary if and only if $AB = BA$ and $A^2 + B^2 = I$.

Solution. Everything comes from expanding T^*T and TT^* for $T = A + iB$. Since $A^* = A$ and $B^* = B$, 7.5(a) and 7.5(b) give $T^* = A - iB$, whence (using $(-i)(i) = 1$)

$$\begin{aligned} T^*T &= (A^2 + B^2) + i(AB - BA), \\ TT^* &= (A^2 + B^2) - i(AB - BA). \end{aligned}$$

(i) If $AB = BA$ and $A^2 + B^2 = I$, the first line gives $T^*T = I$, so T is an isometry by 7.49 ((b) implies (a)); an isometry of V is injective, hence invertible (3.65), hence unitary (7.51).
(ii) If T is unitary then $T^*T = TT^* = I$ by 7.53 ((a) implies (b)), so

$$(A^2 + B^2) + i(AB - BA) = I = (A^2 + B^2) - i(AB - BA).$$

Adding gives $A^2 + B^2 = I$; subtracting gives $2i(AB - BA) = 0$, so $AB = BA$.

Problem 7D.5. Suppose $S \in \mathcal{L}(V)$. Prove that the following are equivalent.
(a) S is a self-adjoint unitary operator.
(b) $S = 2P - I$ for some orthogonal projection P on V.

(c) There exists a subspace U of V such that $Su = u$ for every $u \in U$ and $Sw = -w$ for every $w \in U^\perp$.

Solution. We prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).
(a) $\Rightarrow$ (b). Here $S^2 = S^*S = I$, by self-adjointness and 7.53 ((a) implies (b)). Put $P = \frac{S+I}{2}$; then $P^* = P$ by 7.5(a), (b), (e), and

$$P^2 = \frac{S^2+2S+I}{4} = \frac{2S+2I}{4} = P.$$

Let $U = \operatorname{range} P$. For $v \in V$ write $v = Pv + (v - Pv)$: the first term lies in U, and for $u = Pw \in U$,

$$\langle v - Pv,\, Pw\rangle = \langle Pv - P^2v,\, w\rangle = 0$$

using $P^* = P$ and $P^2 = P$, so $v - Pv \in U^\perp$. This is the decomposition from $V = U \oplus U^\perp$ (6.49), so $P = P_U$ by 6.55, and $2P - I = S$.
(b) $\Rightarrow$ (c). Take U with $P = P_U$. Then $Su = 2u - u = u$ for $u \in U$ by 6.57(b), and $Sw = 0 - w = -w$ for $w \in U^\perp$ by 6.57(c).
(c) $\Rightarrow$ (a). By 6.49 each $v \in V$ is uniquely $v = u + w$ with $u \in U$, $w \in U^\perp$, and $Sv = u - w$. Since $u \perp w$, the Pythagorean theorem (6.12) gives

$$\|Sv\|^2 = \|u\|^2 + \|w\|^2 = \|v\|^2,$$

so S is an isometry; and $S^2v = S(u - w) = u + w = v$ (as $-w \in U^\perp$), so S is invertible and hence unitary (7.51). Finally, writing $v' = u' + w'$ likewise and using $\langle u, w'\rangle = \langle w, u'\rangle = 0$,

$$\langle Sv, v'\rangle = \langle u, u'\rangle - \langle w, w'\rangle = \langle v, Sv'\rangle,$$

so $S^* = S$ by 7.1.

Problem 7D.6. Suppose T_1, T_2 are both normal operators on $\mathbf{F}^3$ with $2, 5, 7$ as eigenvalues. Prove that there exists a unitary operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T_1 = S^*T_2S$.

Solution. Take S to be the operator carrying an eigenbasis of T_1 to the matching eigenbasis of T_2. Set $\lambda_1 = 2$, $\lambda_2 = 5$, $\lambda_3 = 7$, and let T be either T_1 or T_2. Normalizing eigenvectors of T for $\lambda_1, \lambda_2, \lambda_3$ gives unit vectors g_1, g_2, g_3 with $Tg_k = \lambda_k g_k$; these are pairwise orthogonal by 7.22 (T is normal and the λ_k are distinct), hence an orthonormal list of length $3 = \dim \mathbf{F}^3$ and so an orthonormal basis (6.28). No spectral theorem is needed, so this works over $\mathbf{R}$ and over $\mathbf{C}$.
This gives orthonormal bases e_1, e_2, e_3 and f_1, f_2, f_3 with $T_1e_k = \lambda_k e_k$ and $T_2f_k = \lambda_k f_k$. Let $Se_k = f_k$ (3.4); S carries an orthonormal basis to an orthonormal basis, so S is unitary by 7.53 ((d) implies (a)), and $S^* = S^{-1}$ by 7.53 ((a) implies (c)). Hence for each k,

$$S^*T_2Se_k = \lambda_k S^{-1}f_k = \lambda_k e_k = T_1e_k,$$

and the two operators agree on a basis, so $T_1 = S^*T_2S$.

Problem 7D.7. Give an example of two self-adjoint operators $T_1, T_2 \in \mathcal{L}(\mathbf{F}^4)$ such that the eigenvalues of both operators are $2, 5, 7$ but there does not exist a unitary operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $T_1 = S^*T_2S$. Be sure to explain why there is no unitary operator with the required property.

Solution. Take

$$T_1(z_1, z_2, z_3, z_4) = (2z_1,\, 2z_2,\, 5z_3,\, 7z_4),$$
$$T_2(z_1, z_2, z_3, z_4) = (2z_1,\, 5z_2,\, 5z_3,\, 7z_4),$$

whose eigenvalue 2 has multiplicity 2 for T_1 but 1 for T_2.

Each T_j has, with respect to the orthonormal standard basis $e_1, \dots, e_4$, a real diagonal matrix, equal to its own conjugate transpose, so $T_j^* = T_j$ by 7.9. The eigenvalues of T_j are exactly $2, 5, 7$: each occurs on the diagonal, and for $\lambda \notin \{2, 5, 7\}$ the operator $T_j - \lambda I$ sends each e_k to a nonzero multiple of itself, hence is surjective and so invertible (3.65). Reading off the null spaces,

$$\dim \operatorname{null}(T_1 - 2I) = 2, \qquad \dim \operatorname{null}(T_2 - 2I) = 1.$$

Suppose S were unitary with $T_1 = S^* T_2 S$. Then $S^* = S^{-1}$ by 7.53 ((a) implies (c)), so

$$T_1 - 2I = S^{-1}(T_2 - 2I)S.$$

If $(T_1 - 2I)v = 0$ then, applying S, $(T_2 - 2I)Sv = 0$; thus S restricts to an injective map $\operatorname{null}(T_1 - 2I) \to \operatorname{null}(T_2 - 2I)$, forcing $2 \leq 1$ by 3.21, a contradiction.

Problem 7D.8. Prove or give a counterexample: If $S \in \mathcal{L}(V)$ and there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\|Se_k\| = 1$ for each e_k, then S is a unitary operator.

Solution. False: on $V = \mathbf{F}^2$ with the standard (orthonormal) basis, take

$$S(x_1, x_2) = (x_1 + x_2,\ 0).$$

Then $Se_1 = Se_2 = (1, 0)$, so $\|Se_1\| = \|Se_2\| = 1$, yet $S(1, -1) = 0$ with $(1, -1) \neq 0$, so S is not injective, hence not invertible, hence not unitary (7.51).

Problem 7D.9. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Suppose every eigenvalue of T has absolute value 1 and $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that T is a unitary operator.

Solution. A Schur basis for T is forced to be diagonal. Since $\mathbf{F} = \mathbf{C}$, Schur's theorem (6.38) gives an orthonormal basis $e_1, \dots, e_n$ in which $A = \mathcal{M}(T)$ is upper triangular; by 5.41 the diagonal entries are the eigenvalues, so $|A_{k,k}| = 1$ for each k.
Fix k. Upper triangularity gives $Te_k = A_{1,k}e_1 + \cdots + A_{k,k}e_k$, so by 6.30(b) and the hypothesis at $v = e_k$,

$$1 + \sum_{j=1}^{k-1} |A_{j,k}|^2 = \|Te_k\|^2 \leq \|e_k\|^2 = 1.$$

Every term of a nonnegative sum that is at most 0 vanishes, so $A_{j,k} = 0$ for $j < k$. As k was arbitrary and A is already lower-zero, A is diagonal: $Te_k = \lambda_k e_k$ with $|\lambda_k| = 1$.
Thus V has an orthonormal basis of eigenvectors of T with eigenvalues of absolute value 1, so T is unitary by 7.55 ((b) implies (a)).

Problem 7D.10. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is a self-adjoint operator such that $\|Tv\| \leq \|v\|$ for all $v \in V$.
(a) Show that $I - T^2$ is a positive operator.
(b) Show that $T + i\sqrt{I - T^2}$ is a unitary operator.

Solution. (a) $I - T^2$ is self-adjoint, since $(T^2)^* = (T^*)^2 = T^2$ by 7.5(d) and $T^* = T$, so $(I - T^2)^* = I - T^2$ by 7.5(a), (b), (e). And for every $v \in V$,

$$\big\langle (I - T^2)v, v \big\rangle = \|v\|^2 - \langle Tv, T^*v \rangle = \|v\|^2 - \|Tv\|^2 \geq 0$$

by $T^* = T$ and the hypothesis. Hence $I - T^2$ is positive (7.34).
(b) On a spectral basis for T, the operator $T + i\sqrt{I - T^2}$ is diagonal with eigenvalues on the unit circle. Since T is self-adjoint and $\mathbf{F} = \mathbf{C}$, it is normal, so 7.31 gives an orthonormal basis $e_1, \dots, e_n$ with

$Te_k = \lambda_k e_k$, each λ_k real (7.12) and

$$|\lambda_k| = \|Te_k\| \le \|e_k\| = 1,$$

so $1 - \lambda_k^2 \ge 0$. Let $R'e_k = \sqrt{1-\lambda_k^2}\, e_k$: its matrix in this orthonormal basis is real diagonal, so R' is self-adjoint (7.9), and $\langle R'v, v\rangle = \sum_k \sqrt{1-\lambda_k^2}\,|a_k|^2 \ge 0$ for $v = \sum_k a_k e_k$, so R' is positive. Since

$$(R')^2 e_k = (1-\lambda_k^2)e_k = (I - T^2)e_k,$$

uniqueness of positive square roots (7.39) gives $R' = \sqrt{I - T^2} =: R$.
Then $Se_k = (\lambda_k + i\sqrt{1-\lambda_k^2})e_k$ for $S = T + iR$, and $\lambda_k^2 + (1-\lambda_k^2) = 1$, so each eigenvalue has absolute value 1. Thus S is unitary by 7.55 ((b) implies (a)).

Problem 7D.11. Suppose $S \in \mathcal{L}(V)$. Prove that S is a unitary operator if and only if

$$\{Sv : v \in V \text{ and } \|v\| \le 1\} = \{v \in V : \|v\| \le 1\}.$$

Solution. Write $B = \{v \in V : \|v\| \le 1\}$; the condition reads $S(B) = B$.
(i) If S is unitary, then $v \in B$ gives $\|Sv\| = \|v\| \le 1$, so $S(B) \subseteq B$. Conversely $S^{-1} = S^*$ is unitary by 7.53 ((a) implies (c) and (f)), so $w \in B$ gives $\|S^{-1}w\| = \|w\| \le 1$, i.e. $w = S(S^{-1}w) \in S(B)$.
(ii) Suppose $S(B) = B$. Then range $S \supseteq B$ contains every $v \ne 0$ (as $v = \|v\| \cdot v/\|v\|$ and ranges are subspaces), so S is surjective, hence invertible (3.65). Moreover $v/\|v\| \in B$ forces $\|S(v/\|v\|)\| \le 1$, i.e.

$$\|Sv\| \le \|v\| \qquad (v \in V),$$

by homogeneity (the case $v = 0$ being clear). Applying S^{-1} to $S(B) = B$ gives $S^{-1}(B) = B$, so the same homogeneity argument yields $\|S^{-1}w\| \le \|w\|$ for all w; at $w = Sv$ this reads $\|v\| \le \|Sv\|$. Hence $\|Sv\| = \|v\|$ and S is a unitary operator (7.51).

Problem 7D.12. Prove or give a counterexample: If $S \in \mathcal{L}(V)$ is invertible and $\|S^{-1}v\| = \|Sv\|$ for every $v \in V$, then S is unitary.

Solution. False: on $V = \mathbf{F}^2$ take

$$S(x_1, x_2) = \left(2x_2,\ \tfrac{1}{2}x_1\right).$$

Then $S^2 = I$ (Check!), so S is invertible with $S^{-1} = S$ and the hypothesis $\|S^{-1}v\| = \|Sv\|$ holds trivially. But $S(1,0) = (0, \frac{1}{2})$ has norm $\frac{1}{2} \ne 1$, so S is not an isometry and hence not unitary (7.51).

Problem 7D.13. Explain why the columns of a square matrix of complex numbers form an orthonormal list in $\mathbf{C}^n$ if and only if the rows of the matrix form an orthonormal list in $\mathbf{C}^n$.

Solution. Both conditions say Q is unitary, because $Q^*Q = I$ if and only if $QQ^* = I$ for square Q.
Let Q be n-by-n and Q^* its conjugate transpose, $(Q^*)_{j,k} = \overline{Q_{k,j}}$. With the Euclidean inner product on $\mathbf{C}^n$, for all j, k, r,

$$\langle \text{col } k, \text{col } r\rangle = \sum_{j=1}^{n} (Q^*)_{r,j}\, Q_{j,k} = (Q^*Q)_{r,k},$$

$$\langle \text{row } j, \text{row } k\rangle = \sum_{r=1}^{n} Q_{j,r}\, (Q^*)_{r,k} = (QQ^*)_{j,k},$$

so the columns are orthonormal iff $Q^*Q = I$, and the rows are orthonormal iff $QQ^* = I$.
Now let $S \in \mathcal{L}(\mathbf{C}^n)$ have matrix Q with respect to the standard (orthonormal) basis, so $\mathcal{M}(S^*) = Q^*$ by 7.9. If $Q^*Q = I$ then $S^*S = I$, so S is injective ($Sv = 0$ gives $v = S^*Sv = 0$) and hence invertible

(3.65); multiplying $S^*S = I$ on the right by S^{-1} gives $S^* = S^{-1}$, so $SS^* = I$, i.e. $QQ^* = I$. The converse follows by applying this to Q^*, since $(Q^*)^* = Q$.

Problem 7D.14. Suppose $v \in V$ with $\|v\| = 1$ and $b \in \mathbf{F}$. Also suppose $\dim V \geq 2$. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $\langle Sv, v\rangle = b$ if and only if $|b| \leq 1$.

Solution. If such a unitary S exists then $\|Sv\| = \|v\| = 1$, so Cauchy–Schwarz (6.14) gives

$$|b| = |\langle Sv, v\rangle| \leq \|Sv\|\,\|v\| = 1.$$

Conversely, suppose $|b| \leq 1$ and set $c = \sqrt{1 - |b|^2} \in \mathbf{R}$, so that $|b|^2 + c^2 = 1$. Since $\|v\| = 1$, the list v is orthonormal, so by 6.36 it extends to an orthonormal basis $e_1, \dots, e_n$ of V with $e_1 = v$; the hypothesis $n = \dim V \geq 2$ supplies e_2. Define $S \in \mathcal{L}(V)$ on this basis (3.4) by

$$\begin{aligned} Se_1 &= be_1 + ce_2, \qquad Se_2 = -ce_1 + \overline{b}e_2, \\ Se_k &= e_k \quad (k \geq 3). \end{aligned}$$

Then $\|Se_1\|^2 = |b|^2 + c^2 = 1 = c^2 + |b|^2 = \|Se_2\|^2$, while $\overline{c} = c$ gives

$$\langle Se_1, Se_2\rangle = b\,\overline{(-c)} + c\,\overline{\overline{b}} = -bc + cb = 0,$$

and $Se_1, Se_2 \in \operatorname{span}(e_1, e_2)$ is orthogonal to each $Se_k = e_k$ with $k \geq 3$ (Check!). So $Se_1, \dots, Se_n$ is an orthonormal basis of V, whence S is unitary by the equivalence of (d) and (a) in 7.53. Finally

$$\langle Sv, v\rangle = \langle be_1 + ce_2, e_1\rangle = b.$$

Problem 7D.15. Suppose T is a unitary operator on V such that $T - I$ is invertible.
(a) Prove that $(T+I)(T-I)^{-1}$ is a skew operator (meaning that it equals the negative of its adjoint).
(b) Prove that if $\mathbb{F} = \mathbb{C}$, then $i(T+I)(T-I)^{-1}$ is a self-adjoint operator.
[The function $z \mapsto i(z+1)(z-1)^{-1}$ maps the unit circle in $\mathbb{C}$ (except for the point 1) to $\mathbb{R}$. Thus (b) illustrates the analogy between the unitary operators and the unit circle in $\mathbb{C}$, along with the analogy between the self-adjoint operators and $\mathbb{R}$.]

Solution. (a) Write $S = (T+I)(T-I)^{-1}$; we show $S^* = -S$.
Since $(T+I)(T-I) = T^2 - I = (T-I)(T+I)$, multiplying by $(T-I)^{-1}$ on both sides gives

$$S = (T+I)(T-I)^{-1} = (T-I)^{-1}(T+I). \tag{$*$}$$

Because T is unitary, $T^* = T^{-1}$ by the equivalence of (a) and (c) in 7.53, so

$$T^* - I = -T^{-1}(T - I), \qquad T^* + I = T^{-1}(T+I);$$

in particular $T^* - I$ is invertible with $(T^* - I)^{-1} = -(T-I)^{-1}T$. Hence, using the adjoint rules 7.5 together with $(A^*)^{-1} = (A^{-1})^*$ for invertible A (immediate from $(A^{-1})^*A^* = (AA^{-1})^* = I$) applied to $A = T - I$,

$$\begin{aligned} S^* &= \big((T+I)(T-I)^{-1}\big)^* \\ &= \big((T-I)^{-1}\big)^*(T+I)^* \\ &= \big((T-I)^*\big)^{-1}(T^* + I) \\ &= (T^* - I)^{-1}(T^* + I) \\ &= \big(-(T-I)^{-1}T\big)\big(T^{-1}(T+I)\big) \\ &= -(T-I)^{-1}(T+I). \end{aligned}$$

By $(*)$ the last expression equals $-(T+I)(T-I)^{-1} = -S$, so S is skew.

(b) With $\mathbb{F} = \mathbf{C}$, part (a) and $(\lambda B)^* = \overline{\lambda} B^*$ give

$$(iS)^* = \overline{i}\, S^* = (-i)(-S) = iS,$$

so $i(T+I)(T-I)^{-1}$ is self-adjoint.

Problem 7D.16. Suppose $\mathbb{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is self-adjoint. Prove that $(T+iI)(T-iI)^{-1}$ is a unitary operator and 1 is not an eigenvalue of this operator.

Solution. Everything follows from the identity

$$\|(T \pm iI)v\|^2 = \|Tv\|^2 + \|v\|^2 \qquad (v \in V), \tag{$*$}$$

whose cross terms cancel because $\langle Tv, v\rangle = \langle v, Tv\rangle$ for self-adjoint T:

$$\|(T - iI)v\|^2 = \|Tv\|^2 + i\langle Tv, v\rangle - i\langle v, Tv\rangle + \|v\|^2,$$

and likewise with the two cross terms $-i\langle Tv, v\rangle + i\langle v, Tv\rangle$ for $T + iI$.
By $(*)$, $(T \mp iI)v = 0$ forces $\|v\| = 0$, so $T - iI$ and $T + iI$ are injective, hence invertible (3.65); set $S = (T+iI)(T-iI)^{-1}$, a product of invertibles. Given $w \in V$, write $v = (T-iI)^{-1}w$, so $Sw = (T+iI)v$ and $(*)$ gives

$$\|Sw\| = \|(T + iI)v\| = \|(T - iI)v\| = \|w\|.$$

Thus S is an invertible isometry of V, i.e. unitary (7.51).
If $Sw = w$, then with $v = (T-iI)^{-1}w$ we get $(T+iI)v = (T-iI)v$, so $2iv = 0$ and $w = (T-iI)v = 0$. Hence 1 is not an eigenvalue of S.

Problem 7D.17. Explain why the characterizations of unitary matrices given by 7.57 hold.
[7.57: Suppose Q is an n-by-n matrix. Then the following are equivalent.
(a) Q is a unitary matrix.
(b) The rows of Q form an orthonormal list in $\mathbb{F}^n$.
(c) $\|Qv\| = \|v\|$ for every $v \in \mathbb{F}^n$.
(d) $Q^*Q = QQ^* = I$, the n-by-n matrix with 1's on the diagonal and 0's elsewhere.]

Solution. Each of (a)–(d) says that the operator $S \in \mathcal{L}(\mathbb{F}^n)$ with

$$\mathcal{M}\big(S, (e_1, \ldots, e_n)\big) = Q$$

is unitary, where $e_1, \ldots, e_n$ is the standard (orthonormal) basis of $\mathbb{F}^n$ with its Euclidean inner product. Two translations do the work: identifying $v \in \mathbb{F}^n$ with its coordinate column, 3.76 gives $Sv = Qv$, hence $\|Sv\| = \|Qv\|$; and since the basis is orthonormal, 7.9 gives $\mathcal{M}(S^*) = Q^*$, so with 3.43 and the injectivity of $A \mapsto \mathcal{M}(A)$,

$$S^*S = SS^* = I \iff Q^*Q = QQ^* = I.$$

(i) By definition 7.56, (a) says the columns of Q are orthonormal, which by the equivalence of (a) and (e) in 7.49 (both bases taken to be $e_1, \ldots, e_n$) says S is an isometry; an isometry is injective, hence invertible on the finite-dimensional $\mathbb{F}^n$ (3.65), so this is exactly unitarity (7.51).
(ii) (b) says the rows of Q are orthonormal, hence an orthonormal basis (6.28), which is the equivalence of (a) and (e) in 7.53.
(iii) (c) says $\|Sv\| = \|v\|$ for all v by the first translation, i.e. S is an isometry, equivalent to unitarity as in (i).
(iv) (d) says $S^*S = SS^* = I$ by the second translation, which is the equivalence of (a) and (b) in 7.53.

Problem 7D.18. A square matrix A is called *symmetric* if it equals its transpose. Prove that if A is a symmetric matrix with real entries, then there exists a unitary matrix Q with real entries such that Q^*AQ is a diagonal matrix.

Solution. Take Q to be the matrix whose columns are an orthonormal eigenbasis of A acting on $\mathbb{R}^n$. In detail, let $V = \mathbb{R}^n$ with the Euclidean inner product, $e_1, \ldots, e_n$ its standard (orthonormal) basis, and $T \in \mathcal{L}(V)$ the operator with $\mathcal{M}(T, (e_1, \ldots, e_n)) = A$ (legitimate since A has real entries). Orthonormality of the basis lets 7.9 compute

$$\begin{aligned}\mathcal{M}(T^*, (e_1, \ldots, e_n)) &= A^* = A^{\mathrm{t}} = A \\ &= \mathcal{M}(T, (e_1, \ldots, e_n)),\end{aligned}$$

using that A is real and symmetric; hence $T^* = T$. As $\mathbb{F} = \mathbb{R}$ and T is self-adjoint, the real spectral theorem 7.29 supplies an orthonormal basis $f_1, \ldots, f_n$ of $\mathbb{R}^n$ of eigenvectors of T, so $D = \mathcal{M}(T, (f_1, \ldots, f_n))$ is diagonal.
Let Q have k-th column f_k. Its entries are real, its columns are orthonormal, so Q is unitary (7.56); and $Q = \mathcal{M}(I, (f_1, \ldots, f_n), (e_1, \ldots, e_n))$, since the k-th column of the latter is the coordinate column of f_k. The change-of-basis formula 3.84 therefore gives $D = Q^{-1}AQ$, and $Q^{-1} = Q^*$ by 7.57(d), so

$$Q^*AQ = D.$$

Problem 7D.19. Suppose n is a positive integer. For this exercise, we adopt the notation that a typical element z of $\mathbb{C}^n$ is denoted by $z = (z_0, z_1, \ldots, z_{n-1})$. Define linear functionals $\omega_0, \omega_1, \ldots, \omega_{n-1}$ on $\mathbb{C}^n$ by

$$\omega_j(z_0, z_1, \ldots, z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m \, e^{-2\pi i j m/n}.$$

The *discrete Fourier transform* is the operator $\mathcal{F}\colon \mathbb{C}^n \to \mathbb{C}^n$ defined by

$$\mathcal{F}z = \big(\omega_0(z), \omega_1(z), \ldots, \omega_{n-1}(z)\big).$$

(a) Show that $\mathcal{F}$ is a unitary operator on $\mathbb{C}^n$.
(b) Show that if $(z_0, \ldots, z_{n-1}) \in \mathbb{C}^n$ and z_n is defined to equal z_0, then

$$\mathcal{F}^{-1}(z_0, z_1, \ldots, z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, \ldots, z_1).$$

(c) Show that $\mathcal{F}^4 = I$.
[The discrete Fourier transform has many important applications in data analysis. The usual Fourier transform involves expressions of the form $\int_{-\infty}^{\infty} f(x) e^{-2\pi i t x}\, dx$ for complex-valued integrable functions f defined on $\mathbb{R}$.]

Solution. Write $\zeta = e^{2\pi i/n}$, so $e^{-2\pi i j m/n} = \zeta^{-jm}$, $\zeta^n = 1$, and

$$(\mathcal{F}z)_j = \omega_j(z) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m \zeta^{-jm},$$

for $0 \le j \le n-1$, which is linear in z. The engine is the root-of-unity identity: for an integer k,

$$\sum_{j=0}^{n-1} \zeta^{jk} = n \ \text{ if } n \mid k, \quad \text{else } 0,$$

the first case since every term is 1, the second by summing the geometric series with ratio $\zeta^k \neq 1$ to $(1 - (\zeta^n)^k)/(1 - \zeta^k) = 0$.
(a) With $e_0, \ldots, e_{n-1}$ the standard (orthonormal) basis of $\mathbb{C}^n$, the matrix $F = \mathcal{M}(\mathcal{F}, (e_0, \ldots, e_{n-1}))$

has entries $F_{j,m} = \zeta^{-jm}/\sqrt{n}$, so for $m, r \in \{0, \dots, n-1\}$, using $\overline{\zeta^{-jr}} = \zeta^{jr}$,

$$\sum_{j=0}^{n-1} F_{j,m} \overline{F_{j,r}} = \frac{1}{n} \sum_{j=0}^{n-1} \zeta^{j(r-m)} = \begin{cases} 1, & m = r, \\ 0, & m \neq r, \end{cases}$$

because $|r-m| \leq n-1$ makes $n \mid r-m$ equivalent to $r = m$. So the columns of F are orthonormal, hence $\mathcal{F}$ is an isometry by the equivalence of (a) and (e) in 7.49, therefore injective and so invertible (3.65), i.e. unitary (7.51).

(b) Since $\mathcal{F}$ is unitary, $\mathcal{F}^{-1} = \mathcal{F}^*$ by 7.53(c), and 7.9 gives $\mathcal{M}(\mathcal{F}^*) = F^*$ with $(F^*)_{j,m} = \overline{F_{m,j}} = \zeta^{jm}/\sqrt{n}$; thus

$$\left(\mathcal{F}^{-1} z\right)_j = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m \, \zeta^{jm}.$$

Now $w = (z_n, z_{n-1}, \dots, z_1)$ has $w_k = z_{n-k}$ for $0 \leq k \leq n-1$, so substituting $m = n-k$ and using $\zeta^{-jn} = 1$,

$$\begin{aligned} (\mathcal{F} w)_j &= \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} w_k \zeta^{-jk} = \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} z_{n-k} \, \zeta^{-jk} \\ &= \frac{1}{\sqrt{n}} \sum_{m=1}^{n} z_m \, \zeta^{-j(n-m)} = \frac{1}{\sqrt{n}} \sum_{m=1}^{n} z_m \, \zeta^{jm}. \end{aligned}$$

The $m = n$ term is $z_n \zeta^{jn} = z_0 \zeta^{j \cdot 0}$, i.e. the $m = 0$ term, so replacing it leaves the sum unchanged and

$$(\mathcal{F} w)_j = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m \, \zeta^{jm} = \left(\mathcal{F}^{-1} z\right)_j$$

for every j, which is the claimed identity.

(c) Let $R \in \mathcal{L}(\mathbb{C}^n)$ be the reversal $(Rz)_k = z_{(n-k) \bmod n}$, so that $R^2 = I$ (Check!). Since $z_n = z_0$, part (b) reads $\mathcal{F}^{-1} = \mathcal{F} R$; applying $\mathcal{F}^{-1}$ on the left gives $\mathcal{F}^{-2} = R$, and squaring,

$$\mathcal{F}^{-4} = R^2 = I,$$

so $\mathcal{F}^4 = I$.

Problem 7D.20. Suppose A is a square matrix with linearly independent columns. Prove that there exist unique matrices R and Q such that R is lower triangular with only positive numbers on its diagonal, Q is unitary, and $A = RQ$.

Solution. Conjugate-transpose the QR factorization 7.58 of A^*.

Three facts about $M \mapsto M^*$ drive this, all immediate from $(M^*)_{j,k} = \overline{M_{k,j}}$, $(MN)^* = N^* M^*$ and $(M^*)^* = M$: M is lower triangular iff M^* is upper triangular; M has positive diagonal entries iff M^* does; and M is unitary iff M^* is, since 7.57(d) states unitarity as $M^* M = M M^* = I$, a condition symmetric in M and M^*.

First, A^* has linearly independent columns. Indeed, viewing an n-by-n matrix as $\mathcal{M}(T, (e_1, \dots, e_n))$ for the standard basis of $\mathbb{F}^n$, its columns are $\mathcal{M}(Te_1), \dots, \mathcal{M}(Te_n)$, independent exactly when T is invertible, i.e. exactly when the matrix is invertible; and A invertible gives A^* invertible, with $(A^*)^{-1} = (A^{-1})^*$.

Existence: 7.58 applied to A^* yields $A^* = \widetilde{Q} \widetilde{R}$ with $\widetilde{Q}$ unitary and $\widetilde{R}$ upper triangular with positive diagonal. Conjugate-transposing,

$$\begin{aligned} A &= \widetilde{R}^* \, \widetilde{Q}^* = RQ, \\ R &= \widetilde{R}^*, \qquad Q = \widetilde{Q}^*, \end{aligned}$$

and by the three facts R is lower triangular with positive diagonal and Q is unitary.

Uniqueness: if $A = RQ = R'Q'$ with both pairs as stated, then

$$A^* = Q^*R^* = (Q')^*(R')^*$$

are two QR factorizations of A^* (again by the three facts), so 7.58 forces $Q^* = (Q')^*$ and $R^* = (R')^*$, hence $Q = Q'$ and $R = R'$.

7.5 Exercises 7E

Problem 7E.1. Suppose $T \in \mathcal{L}(V, W)$. Show that $T = 0$ if and only if all singular values of T are 0.

Solution. By 7.68(b) the number of positive singular values of T equals $\dim \operatorname{range} T$, so all singular values of T are 0 if and only if $\dim \operatorname{range} T = 0$, i.e. if and only if $T = 0$.
Method (2): if $T = 0$ then $T^*T = 0$, whose only eigenvalue is 0, so by 7.65 every singular value is 0. Conversely, if every singular value is 0 then 0 is the only eigenvalue of T^*T; since T^*T is positive, hence self-adjoint (7.64(a)), the spectral theorem (7.29 or 7.31) gives an orthonormal basis of V of eigenvectors of T^*T, all with eigenvalue 0, so $T^*T = 0$ and therefore $\operatorname{null} T = \operatorname{null} T^*T = V$ by 7.64(b).

Problem 7E.2. Suppose $T \in \mathcal{L}(V, W)$ and $s > 0$. Prove that s is a singular value of T if and only if there exist nonzero vectors $v \in V$ and $w \in W$ such that

$$Tv = sw \quad \text{and} \quad T^*w = sv.$$

[The vectors v, w satisfying both equations above are called a Schmidt pair. Erhard Schmidt introduced the concept of singular values in 1907.]

Solution. Since $s > 0$, definition 7.65 makes s a singular value of T exactly when s^2 is an eigenvalue of T^*T.
If s is a singular value, pick $v \neq 0$ with $T^*Tv = s^2v$ and set $w = Tv/s$. Then $Tv = sw$ and

$$T^*w = \tfrac{1}{s}T^*Tv = \tfrac{1}{s}\,s^2v = sv,$$

while $\|Tv\|^2 = \langle T^*Tv, v\rangle = s^2\|v\|^2 > 0$ gives $w \neq 0$.
Conversely, if v, w are nonzero with $Tv = sw$ and $T^*w = sv$, then

$$T^*Tv = T^*(sw) = s(sv) = s^2v,$$

so s^2 is an eigenvalue of T^*T and s is a singular value of T.

Problem 7E.3. Give an example of $T \in \mathcal{L}(\mathbf{C}^2)$ such that 0 is the only eigenvalue of T and the singular values of T are $5, 0$.

Solution. Take $T(z_1, z_2) = (5z_2, 0)$, with matrix $\begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix}$ in the standard (orthonormal) basis.
$T^2 = 0$, so 0 is the only eigenvalue of T, and it is an eigenvalue since $T(1,0) = 0$. Because the basis is orthonormal, 7.9 makes $\mathcal{M}(T^*)$ the conjugate transpose, so

$$\mathcal{M}(T^*T) = \begin{pmatrix} 0 & 0 \\ 5 & 0 \end{pmatrix}\begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix},$$

whose eigenvalues 25 and 0 each have one-dimensional eigenspace. Hence the singular values of T are $5, 0$.

Problem 7E.4. Suppose that $T \in \mathcal{L}(V, W)$, s_1 is the largest singular value of T, and s_n is the smallest singular value of T. Prove that

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = [s_n, s_1].$$

Solution. Everything follows from the formula $\|Tv\|^2 = \sum_k s_k^2 |a_k|^2$ in a suitable orthonormal eigenbasis. Let $n = \dim V \geq 1$ and $s_1 \geq \cdots \geq s_n \geq 0$ be the singular values. Since T^*T is positive, hence self-adjoint (7.64(a)), the spectral theorem (7.29 or 7.31) gives an orthonormal basis $e_1, \ldots, e_n$ of V with $T^*Te_k = s_k^2 e_k$ (concatenate orthonormal bases of the eigenspaces, as in the proof of 7.70; multiplicities match 7.65 by construction). For $v = \sum_k a_k e_k$ with $\|v\| = 1$, so $\sum_k |a_k|^2 = 1$ by 6.30(b),

$$\|Tv\|^2 = \langle T^*Tv, v\rangle = \sum_{k=1}^{n} s_k^2 |a_k|^2,$$

which lies between $s_n^2 \sum_k |a_k|^2 = s_n^2$ and $s_1^2 \sum_k |a_k|^2 = s_1^2$; taking square roots gives $\|Tv\| \in [s_n, s_1]$. Conversely, let $c \in [s_n, s_1]$. If $s_1 = s_n$ then $v = e_1$ gives $\|Tv\| = s_1 = c$. If $s_1 > s_n$, put $t = (c^2 - s_n^2)/(s_1^2 - s_n^2) \in [0, 1]$ and $v = \sqrt{t}\, e_1 + \sqrt{1-t}\, e_n$, a unit vector with

$$\|Tv\|^2 = s_1^2 t + s_n^2(1-t) = s_n^2 + (c^2 - s_n^2) = c^2,$$

so $\|Tv\| = c$.

Problem 7E.5. Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is defined by $T(x, y) = (-4y, x)$. Find the singular values of T.

Solution. The singular values of T are $4, 1$.
Indeed, in the standard (orthonormal) basis $\mathcal{M}(T) = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix}$, and 7.9 makes $\mathcal{M}(T^*)$ its conjugate transpose, so

$$\mathcal{M}(T^*T) = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 16 \end{pmatrix},$$

whose eigenvalues 16 and 1 each have one-dimensional eigenspace.

Problem 7E.6. Find the singular values of the differentiation operator $D \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ defined by $Dp = p'$, where the inner product on $\mathcal{P}_2(\mathbf{R})$ is as in Example 6.34.

Solution. The singular values of D are $\sqrt{15}, \sqrt{3}, 0$.
Example 6.34 supplies the orthonormal basis

$$e_1 = \sqrt{\tfrac{1}{2}}, \qquad e_2 = \sqrt{\tfrac{3}{2}}\, x,$$
$$e_3 = \sqrt{\tfrac{45}{8}}\left(x^2 - \tfrac{1}{3}\right)$$

for the inner product $\langle p, q\rangle = \int_{-1}^{1} pq$. Using $1 = \sqrt{2}\, e_1$ and $x = \sqrt{2/3}\, e_2$,

$$De_1 = 0, \qquad De_2 = \sqrt{\tfrac{3}{2}} = \sqrt{3}\, e_1,$$
$$De_3 = 2\sqrt{\tfrac{45}{8}}\, x = \sqrt{15}\, e_2,$$

so $\mathcal{M}(D)$ is strictly upper triangular with superdiagonal $\sqrt{3}, \sqrt{15}$. Since the basis is orthonormal, 7.9 gives $\mathcal{M}(D^*) = \mathcal{M}(D)^{\mathrm{t}}$, whence

$$\mathcal{M}(D^*D) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 15 \end{pmatrix},$$

with eigenvalues $15, 3, 0$, each eigenspace one-dimensional.

Problem 7E.7. Suppose that $T \in \mathcal{L}(V)$ is self-adjoint or that $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, each included in this list as many times as the dimension of the corresponding eigenspace. Show that the singular values of T are $|\lambda_1|, \ldots, |\lambda_n|$, after these numbers have been sorted into decreasing order.

Solution. The point is that $T^*Te_k = |\lambda_k|^2 e_k$ on an orthonormal eigenbasis of T.
In either hypothesis T is normal and the spectral theorem applies (7.29 when $\mathbf{F} = \mathbf{R}$ and T is self-adjoint; 7.31 when $\mathbf{F} = \mathbf{C}$), so T is diagonalizable and 5.55(c) writes V as the direct sum of its eigenspaces; concatenating orthonormal bases of these (orthogonal to one another by 7.22) gives an orthonormal basis $e_1, \ldots, e_n$ of V with $Te_k = \lambda_k e_k$, where $\lambda_1, \ldots, \lambda_n$ is the list in the statement up to order (immaterial, as the conclusion sorts). Since T is normal, 7.21(e) gives $T^*e_k = \overline{\lambda_k} e_k$, so

$$T^*Te_k = \lambda_k \overline{\lambda_k}\, e_k = |\lambda_k|^2 e_k.$$

Expanding $v = \sum_k a_k e_k$ shows $T^*Tv = \mu v$ iff $a_k = 0$ whenever $|\lambda_k|^2 \neq \mu$, so

$$E(\mu, T^*T) = \operatorname{span}\{e_k : |\lambda_k|^2 = \mu\},$$

whence the eigenvalues of T^*T are the distinct $|\lambda_k|^2$, each with eigenspace dimension $\#\{k : |\lambda_k| = \sqrt{\mu}\}$. Taking nonnegative square roots in 7.65 therefore yields exactly $|\lambda_1|, \ldots, |\lambda_n|$ sorted into decreasing order.

Problem 7E.8. Suppose $T \in \mathcal{L}(V, W)$. Suppose $s_1 \geq s_2 \geq \cdots \geq s_m > 0$ and $e_1, \ldots, e_m$ is an orthonormal list in V and $f_1, \ldots, f_m$ is an orthonormal list in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$.
(a) Prove that $f_1, \ldots, f_m$ is an orthonormal basis of $\operatorname{range} T$.
(b) Prove that $e_1, \ldots, e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$.
(c) Prove that $s_1, \ldots, s_m$ are the positive singular values of T.
(d) Prove that if $k \in \{1, \ldots, m\}$, then e_k is an eigenvector of T^*T with corresponding eigenvalue s_k^2.
(e) Prove that

$$TT^*w = s_1^2 \langle w, f_1 \rangle f_1 + \cdots + s_m^2 \langle w, f_m \rangle f_m$$

for all $w \in W$.

Solution. Two formulas carry all five parts: $Te_j = s_j f_j$ (take $v = e_j$ above) and

$$T^*w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m, \tag{$*$}$$

the latter because the right side Sw satisfies $\langle Tv, w \rangle = \langle v, Sw \rangle$ for all v, w:

$$\begin{aligned} \Big\langle \sum_k s_k \langle v, e_k \rangle f_k, w \Big\rangle &= \sum_k s_k \langle v, e_k \rangle \langle f_k, w \rangle \\ &= \Big\langle v, \sum_k s_k \langle w, f_k \rangle e_k \Big\rangle, \end{aligned}$$

and the adjoint is unique. Consequently, by orthonormality of $f_1, \ldots, f_m$,

$$T^*Te_k = s_k T^* f_k = s_k^2 e_k \qquad (1 \leq k \leq m). \tag{$\dagger$}$$

Write $E = \operatorname{span}(e_1, \ldots, e_m)$, $F = \operatorname{span}(f_1, \ldots, f_m)$, $n = \dim V$.
(a) The formula for T gives $\operatorname{range} T \subseteq F$, and $f_j = Te_j / s_j \in \operatorname{range} T$ (legitimate as $s_j > 0$) gives the reverse. So $f_1, \ldots, f_m$ is an orthonormal, hence linearly independent, spanning list of $\operatorname{range} T$.
(b) Since $f_1, \ldots, f_m$ is linearly independent and every $s_k > 0$,

$$Tv = 0 \iff \langle v, e_k \rangle = 0 \text{ for all } k \iff v \in E^\perp,$$

the second equivalence by expanding any $u \in E$ in $e_1, \ldots, e_m$. So $\operatorname{null} T = E^\perp$, and 6.52 gives $(\operatorname{null} T)^\perp = (E^\perp)^\perp = E$.
(c) Extend $e_1, \ldots, e_m$ by an orthonormal basis of $E^\perp$ (6.35, using $V = E \oplus E^\perp$ from 6.49) to an orthonormal basis $e_1, \ldots, e_n$ of V. For $k > m$, part (b) gives $e_k \in \operatorname{null} T$, so $T^*Te_k = 0$; with (†) this reads $T^*Te_k = d_k e_k$ where $d_k = s_k^2$ for $k \leq m$ and $d_k = 0$ otherwise, a decreasing list. Expanding $v = \sum_k a_k e_k$ shows $E(\lambda, T^*T) = \operatorname{span}\{e_k : d_k = \lambda\}$, so each eigenvalue of T^*T occurs in $d_1, \ldots, d_n$ exactly $\dim E(\lambda, T^*T)$ times. Taking nonnegative square roots, 7.65 gives the singular values $s_1, \ldots, s_m, 0, \ldots, 0$, whose positive members are $s_1, \ldots, s_m$.
(d) This is (†), with $e_k \neq 0$ since $\|e_k\| = 1$.
(e) Apply T to (∗) and use $Te_k = s_k f_k$:

$$TT^*w = \sum_{k=1}^{m} s_k \langle w, f_k \rangle Te_k = \sum_{k=1}^{m} s_k^2 \langle w, f_k \rangle f_k.$$

Problem 7E.9. Suppose $T \in \mathcal{L}(V, W)$. Show that T and T^* have the same positive singular values.

Solution. The singular value decomposition of T^* reuses the data of the one for T.
If T has no positive singular values then $\dim \operatorname{range} T = 0$ by 7.68(b), so $T = T^* = 0$ and neither does T^*. Otherwise let $s_1 \geq \cdots \geq s_m > 0$ be the positive singular values of T; by 7.70 there are orthonormal lists $e_1, \ldots, e_m$ in V and $f_1, \ldots, f_m$ in W with

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m,$$

and then 7.75 (see 7.77) gives

$$T^*w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m.$$

The latter is exactly the hypothesis of 7E.8 for $T^* \in \mathcal{L}(W, V)$, with the f's in the domain and the e's in the codomain, so 7E.8(c) makes $s_1, \ldots, s_m$ the positive singular values of T^*.

Problem 7E.10. Suppose $T \in \mathcal{L}(V, W)$ has singular values $s_1, \ldots, s_n$. Prove that if T is an invertible linear map, then T^{-1} has singular values

$$\frac{1}{s_n}, \ldots, \frac{1}{s_1}.$$

Solution. Invert the singular value decomposition term by term and reindex.
Here $n = \dim V$ and $s_1 \geq \cdots \geq s_n \geq 0$ by 7.65. Since T is invertible it is injective, so $s_n > 0$ by 7.68(a), and surjective, so $\dim W = n$. All n singular values being positive, 7.70 provides orthonormal lists $e_1, \ldots, e_n$ in V and $f_1, \ldots, f_n$ in W with

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n;$$

having length $n = \dim V = \dim W$, these are in fact orthonormal bases, and $Te_k = s_k f_k$. Define $S \in \mathcal{L}(W, V)$ by $Sw = \sum_k s_k^{-1} \langle w, f_k \rangle e_k$. Then

$$S(Te_k) = S(s_k f_k) = e_k, \qquad T(Sf_k) = T(s_k^{-1} e_k) = f_k,$$

so ST and TS agree with the identity on bases, giving $T^{-1} = S$. Reindexing by $k \mapsto n + 1 - k$ makes the coefficients $t_k = 1/s_{n+1-k}$ decreasing and positive, with f_{n+1-k} orthonormal in W and e_{n+1-k} orthonormal in V:

$$T^{-1}w = t_1 \langle w, f_n \rangle e_n + \cdots + t_n \langle w, f_1 \rangle e_1.$$

This is the hypothesis of 7E.8 for T^{-1}, so by 7E.8(c) the positive singular values of T^{-1} are $1/s_n, \ldots, 1/s_1$; since its full list has length $\dim W = n$, no zeros remain.

Problem 7E.11. Suppose that $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_n$ is an orthonormal basis of V. Let $s_1, \dots, s_n$ denote the singular values of T.
(a) Prove that $\|Tv_1\|^2 + \cdots + \|Tv_n\|^2 = s_1^2 + \cdots + s_n^2$.
(b) Prove that if $W = V$ and T is a positive operator, then

$$\langle Tv_1, v_1 \rangle + \cdots + \langle Tv_n, v_n \rangle = s_1 + \cdots + s_n.$$

See the comment after Exercise 5 in Section 7A.

Solution. Both parts reduce to Parseval's identity 6.30(b), which for fixed $u \in V$ gives $\sum_i |\langle v_i, u \rangle|^2 = \sum_i |\langle u, v_i \rangle|^2 = \|u\|^2$.
(a) Let $s_1 \geq \cdots \geq s_m > 0$ be the positive singular values (if $m = 0$ then $T = 0$ by 7.68(b) and both sides vanish), and take the decomposition 7.70,

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m,$$

with $e_1, \dots, e_m$ orthonormal in V and $f_1, \dots, f_m$ orthonormal in W. By 6.24, $\|Tv\|^2 = \sum_{k \leq m} s_k^2 |\langle v, e_k \rangle|^2$, so summing over $v = v_i$ and applying Parseval to each e_k,

$$\begin{aligned}\sum_{i=1}^n \|Tv_i\|^2 &= \sum_{k=1}^m s_k^2 \sum_{i=1}^n |\langle v_i, e_k \rangle|^2 \\ &= \sum_{k=1}^m s_k^2 \|e_k\|^2 = s_1^2 + \cdots + s_n^2,\end{aligned}$$

the last step because $s_k = 0$ for $k > m$.
(b) Since T is positive it is self-adjoint with nonnegative eigenvalues (7.38), so the spectral theorem (7.29 or 7.31) gives an orthonormal basis $u_1, \dots, u_n$ of V with $Tu_k = \lambda_k u_k$, relabelled so $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Then $T^*T = T^2$ satisfies $T^2 u_k = \lambda_k^2 u_k$, and expanding $v = \sum_k a_k u_k$ gives $E(\lambda, T^2) = \operatorname{span}\{u_k : \lambda_k^2 = \lambda\}$, so the eigenvalues of T^*T listed with multiplicity are $\lambda_1^2, \dots, \lambda_n^2$ and 7.65 yields $s_k = \lambda_k$. Now by 6.30(a) and orthonormality of the u_k,

$$\begin{aligned}\langle Tv_i, v_i \rangle &= \Big\langle \sum_k \lambda_k \langle v_i, u_k \rangle u_k, \sum_j \langle v_i, u_j \rangle u_j \Big\rangle \\ &= \sum_{k=1}^n \lambda_k |\langle v_i, u_k \rangle|^2,\end{aligned}$$

so summing over i and applying Parseval to each u_k,

$$\sum_{i=1}^n \langle Tv_i, v_i \rangle = \sum_{k=1}^n \lambda_k \|u_k\|^2 = s_1 + \cdots + s_n.$$

Problem 7E.12. (a) Give an example of a finite-dimensional vector space and an operator T on it such that the singular values of T^2 do not equal the squares of the singular values of T.
(b) Suppose $T \in \mathcal{L}(V)$ is normal. Prove that the singular values of T^2 equal the squares of the singular values of T.

Solution. Both parts use the diagonal counting fact: if $Au_k = d_k u_k$ on an orthonormal basis $u_1, \dots, u_n$ of V with $d_1 \geq \cdots \geq d_n$, then expanding $v = \sum_k a_k u_k$ gives $E(\lambda, A) = \operatorname{span}\{u_k : d_k = \lambda\}$, so the eigenvalues of A listed decreasingly with multiplicity are exactly $d_1, \dots, d_n$.
(a) Take $T \in \mathcal{L}(\mathbf{F}^2)$ with $T(x, y) = (y, 0)$. Then $T^2 = 0$, so the singular values of T^2 are $0, 0$. But $\langle T(x, y), (a, b) \rangle = y\bar{a} = \langle (x, y), (0, a) \rangle$ gives $T^*(a, b) = (0, a)$, hence

$$T^*T(x, y) = (0, y),$$

so $T^*Te_2 = e_2$ and $T^*Te_1 = 0$: the singular values of T are $1, 0$, whose squares $1, 0$ differ from $0, 0$.
(b) Normality gives $T^*T = TT^*$, so with $P = T^*T$ and $(T^2)^* = (T^*)^2$ (7.5(d)),

$$(T^2)^*T^2 = T^*(T^*T)T = T^*(TT^*)T = P^2.$$

P is positive (7.64(a)), so by 7.38 and the spectral theorem (7.29 or 7.31) there is an orthonormal basis $u_1, \dots, u_n$ with $Pu_k = \mu_k u_k$ and $\mu_1 \geq \cdots \geq \mu_n \geq 0$; then $P^2u_k = \mu_k^2 u_k$ with $\mu_1^2 \geq \cdots \geq \mu_n^2$. By the counting fact and 7.65, the singular values of T are $s_k = \sqrt{\mu_k}$ and those of T^2 are

$$\sqrt{\mu_k^2} = \mu_k = s_k^2,$$

using $\mu_k \geq 0$.

Problem 7E.13. Suppose $T_1, T_2 \in \mathcal{L}(V)$. Prove that T_1 and T_2 have the same singular values if and only if there exist unitary operators $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1T_2S_2$.

Solution. If $T_1 = S_1T_2S_2$ with S_1, S_2 unitary, then $S_1^*S_1 = I$ and $S_2^* = S_2^{-1}$ (7.53), so by 7.5(d)

$$T_1^*T_1 = S_2^*T_2^*S_1^*S_1T_2S_2 = S_2^{-1}(T_2^*T_2)S_2.$$

Then $(T_1^*T_1)v = \lambda v \iff (T_2^*T_2)(S_2v) = \lambda(S_2v)$, so S_2 restricts to an isomorphism $E(\lambda, T_1^*T_1) \to E(\lambda, T_2^*T_2)$; the two operators thus have the same eigenvalues with the same eigenspace dimensions, and 7.65 produces the same list of singular values.
Conversely, let $s_1 \geq \cdots \geq s_m > 0$ be the common positive singular values. By 7.70 there are orthonormal lists in V with

$$T_1v = \sum_{k=1}^{m} s_k\langle v, e_k\rangle f_k, \qquad T_2v = \sum_{k=1}^{m} s_k\langle v, g_k\rangle h_k$$

(when $m = 0$ both maps are 0 and any unitaries work). Extend each of the four lists to an orthonormal basis of V by adjoining an orthonormal basis of the orthogonal complement of its span (6.35, using $V = U \oplus U^\perp$ from 6.49). Let $S_2e_k = g_k$ and $S_1h_k = f_k$ for $1 \leq k \leq n$ (3.4); each carries an orthonormal basis to an orthonormal basis, hence is unitary by the equivalence of (a) and (d) in 7.53. For $v \in V$, 6.30(a) gives $S_2v = \sum_k \langle v, e_k\rangle g_k$, so $\langle S_2v, g_k\rangle = \langle v, e_k\rangle$ and therefore

$$\begin{aligned} S_1T_2S_2v &= S_1\sum_{k=1}^{m} s_k\langle v, e_k\rangle h_k \\ &= \sum_{k=1}^{m} s_k\langle v, e_k\rangle f_k = T_1v. \end{aligned}$$

Problem 7E.14. Suppose $T \in \mathcal{L}(V, W)$. Let s_n denote the smallest singular value of T. Prove that $s_n\|v\| \leq \|Tv\|$ for every $v \in V$.

Solution. This is the lower bound in 7E.4, now for all v rather than unit vectors.
In detail, T^*T is positive (7.64(a)), so the spectral theorem (7.29 or 7.31) with 7.38 gives an orthonormal basis $u_1, \dots, u_n$ of V with $T^*Tu_k = d_ku_k$ and $d_1 \geq \cdots \geq d_n \geq 0$. Expanding $v = \sum_k a_ku_k$ shows $E(\lambda, T^*T) = \operatorname{span}\{u_k : d_k = \lambda\}$, so the eigenvalues of T^*T listed decreasingly with multiplicity are $d_1, \dots, d_n$ and 7.65 gives $d_k = s_k^2$. Hence, for $v = \sum_k a_ku_k$ with $a_k = \langle v, u_k\rangle$ (6.30(a)), using $(T^*)^* = T$

(7.5(c)) and orthonormality,

$$\begin{aligned}\|Tv\|^2 = \langle T^*Tv, v\rangle &= \sum_{k=1}^n s_k^2|a_k|^2 \\ &\geq s_n^2 \sum_{k=1}^n |a_k|^2 = s_n^2\|v\|^2\end{aligned}$$

by 6.24. Taking nonnegative square roots gives $s_n\|v\| \leq \|Tv\|$.

Problem 7E.15. Suppose $T \in \mathcal{L}(V)$ and $s_1 \geq \cdots \geq s_n$ are the singular values of T. Prove that if λ is an eigenvalue of T, then $s_1 \geq |\lambda| \geq s_n$.

Solution. Apply the bounds $s_n\|v\| \leq \|Tv\| \leq s_1\|v\|$ to a unit eigenvector.
Those bounds hold for every $v \in V$: as in 7E.14, T^*T is positive (7.64(a)), so the spectral theorem (7.29 or 7.31) with 7.38 gives an orthonormal basis $e_1, \ldots, e_n$ of V with $T^*Te_k = \mu_k e_k$, $\mu_1 \geq \cdots \geq \mu_n \geq 0$; the counting argument of 7E.14 identifies each μ as occurring $\dim E(\mu, T^*T)$ times in this list, so $s_k = \sqrt{\mu_k}$ by 7.65 and $T^*Te_k = s_k^2 e_k$. Then for $v = \sum_k a_k e_k$, using $(T^*)^* = T$ (7.5(c)) and orthonormality,

$$\|Tv\|^2 = \langle T^*Tv, v\rangle = \sum_{k=1}^n s_k^2|a_k|^2 \in [s_n^2\|v\|^2,\ s_1^2\|v\|^2].$$

Now if $Tv = \lambda v$ with $\|v\| = 1$ (normalize any eigenvector), then $\|Tv\| = |\lambda|$, so $s_n \leq |\lambda| \leq s_1$.

Problem 7E.16. Suppose $T \in \mathcal{L}(V, W)$. Prove that $(T^*)^\dagger = (T^\dagger)^*$.
Compare the result in this exercise to the analogous result for invertible linear maps [see 7.5(f)].

Solution. Both sides equal $v \mapsto \sum_k s_k^{-1}\langle v, e_k\rangle f_k$ in the singular value decomposition of T.
If $T = 0$ then $T^* = 0$ and $T^\dagger = 0$ (6.68, since $(\operatorname{null} T)^\perp = \{0\}$ and $P_{\operatorname{range} T} = 0$), so both sides are 0. Otherwise let $s_1 \geq \cdots \geq s_m > 0$ be the positive singular values (at least one exists, by 7E.1) and take 7.70:

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m,$$

the shape 7.76 required by 7.75, with $e_1, \ldots, e_m$ orthonormal in V and $f_1, \ldots, f_m$ orthonormal in W.
(i) By 7.77, $T^*w = \sum_k s_k\langle w, f_k\rangle e_k$, which is again the shape 7.76 (now with the f's in the domain), so 7E.9 makes $s_1, \ldots, s_m$ the positive singular values of T^* and 7.78 applied to T^* gives

$$(T^*)^\dagger v = \sum_{k=1}^m \frac{\langle v, e_k\rangle}{s_k} f_k.$$

(ii) By 7.78 applied to T, $T^\dagger w = \sum_k s_k^{-1}\langle w, f_k\rangle e_k$, so for $w \in W$, $v \in V$, using $s_k > 0$ real and $\overline{\langle e_k, v\rangle} = \langle v, e_k\rangle$,

$$\begin{aligned}\langle T^\dagger w, v\rangle &= \sum_{k=1}^m \frac{\langle w, f_k\rangle\langle e_k, v\rangle}{s_k} \\ &= \Big\langle w, \sum_{k=1}^m \frac{\langle v, e_k\rangle}{s_k} f_k\Big\rangle,\end{aligned}$$

so by uniqueness of the adjoint (7.1), $(T^\dagger)^*v = \sum_k s_k^{-1}\langle v, e_k\rangle f_k$, the formula in (i).

Problem 7E.17. Suppose $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if $T^\dagger$ is self-adjoint.

Solution. If $T^* = T$, then 7E.16 gives

$$(T^\dagger)^* = (T^*)^\dagger = T^\dagger,$$

so $T^\dagger$ is self-adjoint. Conversely, if $T^\dagger$ is self-adjoint, that implication applied to $T^\dagger$ makes $(T^\dagger)^\dagger$ self-adjoint, and $(T^\dagger)^\dagger = T$ by Exercise 23 in Section 6C.

7.6 Exercises 7F

Problem 7F.1. Prove that if $S, T \in \mathcal{L}(V, W)$, then $\big|\, \|S\| - \|T\| \,\big| \le \|S - T\|$.
The inequality above is called the *reverse triangle inequality.*

Solution. Write $S = (S - T) + T$ and apply the triangle inequality 7.87(d):

$$\|S\| \le \|S - T\| + \|T\|, \qquad \text{so} \qquad \|S\| - \|T\| \le \|S - T\|.$$

Interchanging S and T gives $\|T\| - \|S\| \le \|T - S\| = \|S - T\|$, the last equality by 7.87(c) with $\lambda = -1$. Since $\big|\, \|S\| - \|T\| \,\big|$ is one of these two differences, it is at most $\|S - T\|$.

Problem 7F.2. Suppose that $T \in \mathcal{L}(V)$ is self-adjoint or that $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that

$$\|T\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } T\}.$$

Solution. The spectral theorem (7.29 when $\mathbf{F} = \mathbf{R}$ and T is self-adjoint; 7.31 when $\mathbf{F} = \mathbf{C}$ and T is normal, which covers a self-adjoint T there) gives an orthonormal basis $e_1, \dots, e_n$ of V with $Te_k = \lambda_k e_k$. The eigenvalues of T are exactly $\lambda_1, \dots, \lambda_n$: each λ_k is one, and if $Tv = \lambda v$ with $v = a_1 e_1 + \cdots + a_n e_n \ne 0$, comparing coefficients gives $\lambda_k a_k = \lambda a_k$ for every k, so $\lambda = \lambda_k$ for any k with $a_k \ne 0$. Hence the right side is $M := \max\{|\lambda_1|, \dots, |\lambda_n|\}$, a maximum over a finite nonempty set since $n = \dim V \ge 1$.
If $\|v\| \le 1$, write $v = a_1 e_1 + \cdots + a_n e_n$; then $Tv = \lambda_1 a_1 e_1 + \cdots + \lambda_n a_n e_n$, so two applications of 6.24 give

$$\begin{aligned} \|Tv\|^2 &= |\lambda_1|^2 |a_1|^2 + \cdots + |\lambda_n|^2 |a_n|^2 \\ &\le M^2 \big(|a_1|^2 + \cdots + |a_n|^2\big) = M^2 \|v\|^2 \le M^2. \end{aligned}$$

Thus $\|T\| \le M$ by 7.86. Choosing k with $|\lambda_k| = M$, the unit vector e_k has $\|Te_k\| = |\lambda_k|\, \|e_k\| = M$, so $\|T\| \ge M$. Hence $\|T\| = M$.

Problem 7F.3. Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Prove that

$$\|Tv\| = \|T\|\, \|v\| \iff T^* T v = \|T\|^2 v.$$

Solution. Write $c = \|T\| \ge 0$. The adjoint identity 7.1 with the pair v, Tv gives

$$\|Tv\|^2 = \langle v, T^* T v \rangle = \langle T^* T v, v \rangle,$$

the second equality because the left side is real, so the middle term equals its own conjugate.
($\Leftarrow$) If $T^* T v = c^2 v$, then $\|Tv\|^2 = \langle v, c^2 v \rangle = c^2 \|v\|^2$ (the conjugate of c^2 is c^2, as c is real); taking nonnegative square roots gives $\|Tv\| = \|T\|\, \|v\|$.
($\Rightarrow$) Suppose $\|Tv\| = c\|v\|$. By 7.89 applied to T^* and the vector Tv, then 7.91 and the hypothesis,

$$\|T^* T v\| \le \|T^*\|\, \|Tv\| = c \cdot c\|v\| = c^2 \|v\|,$$

while the identity above with c^2 real gives $\langle T^* T v, c^2 v \rangle = c^2 \|Tv\|^2 = c^4 \|v\|^2$, a real number, so its conjugate $\langle c^2 v, T^*Tv\rangle$ equals $c^4\|v\|^2$ too. Expanding,

$$\begin{aligned}\|T^*Tv - c^2 v\|^2 &= \|T^*Tv\|^2 - 2c^4\|v\|^2 + c^4\|v\|^2 \\ &\le c^4\|v\|^2 - 2c^4\|v\|^2 + c^4\|v\|^2 = 0.\end{aligned}$$

Hence $T^*Tv - c^2 v = 0$, that is, $T^*Tv = \|T\|^2 v$.

Problem 7F.4. Suppose $T \in \mathcal{L}(V, W)$, $v \in V$, and $\|Tv\| = \|T\|\,\|v\|$. Prove that if $u \in V$ and $\langle u, v\rangle = 0$, then $\langle Tu, Tv\rangle = 0$.

Solution. The hypothesis is the left side of the equivalence in Exercise 3 of this section, so $T^*Tv = \|T\|^2 v$. Hence, by the adjoint identity 7.1,

$$\begin{aligned}\langle Tu, Tv\rangle = \langle u, T^*Tv\rangle &= \left\langle u, \|T\|^2 v\right\rangle \\ &= \|T\|^2\langle u, v\rangle = 0,\end{aligned}$$

the conjugate on $\|T\|^2$ being dropped because it is a nonnegative real number.

Problem 7F.5. Suppose U is a finite-dimensional inner product space, $T \in \mathcal{L}(V, U)$, and $S \in \mathcal{L}(U, W)$. Prove that

$$\|ST\| \le \|S\|\,\|T\|.$$

Solution. For $v \in V$, two applications of 7.89 (to S with the vector $Tv \in U$, then to T with v, multiplying by $\|S\| \ge 0$) give

$$\|(ST)v\| = \|S(Tv)\| \le \|S\|\,\|Tv\| \le \|S\|\,\|T\|\,\|v\|.$$

Thus $c = \|S\|\,\|T\|$ is a number with $\|(ST)v\| \le c\|v\|$ for all $v \in V$, and $\|ST\|$ is the smallest such number by 7.88(c). Hence $\|ST\| \le \|S\|\,\|T\|$.

Problem 7F.6. Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\|ST\| = \|TS\|$.

Solution. False. Take $V = \mathbf{F}^2$ with its usual inner product and

$$S(a_1, a_2) = (a_2, 0), \qquad T(a_1, a_2) = (a_1, 0).$$

Then $(ST)(a_1, a_2) = S(a_1, 0) = (0, 0)$, so $ST = 0$ and $\|ST\| = 0$ by 7.87(b). But $(TS)(a_1, a_2) = T(a_2, 0) = (a_2, 0)$, so $\|(TS)v\| \le \|v\|$ for all v, giving $\|TS\| \le 1$ by 7.86, while $\|(TS)e_2\| = \|e_1\| = 1$ with $\|e_2\| = 1$ gives $\|TS\| \ge 1$. Hence $\|ST\| = 0 \ne 1 = \|TS\|$.

Problem 7F.7. Show that defining $d(S, T) = \|S - T\|$ for $S, T \in \mathcal{L}(V, W)$ makes d a metric on $\mathcal{L}(V, W)$.
This exercise is intended for readers who are familiar with metric spaces.

Solution. Each of the four metric axioms is one part of 7.87 applied to $S - T \in \mathcal{L}(V, W)$.
(i) Nonnegativity: $d(S, T) = \|S - T\| \ge 0$ by 7.87(a).
(ii) Definiteness: by 7.87(b),

$$d(S, T) = 0 \iff \|S - T\| = 0 \iff S - T = 0 \iff S = T.$$

(iii) Symmetry: since $T - S = (-1)(S - T)$, property 7.87(c) with $\lambda = -1$ gives

$$d(T, S) = \|T - S\| = |-1|\,\|S - T\| = d(S, T).$$

(iv) Triangle inequality: since $S - U = (S - T) + (T - U)$, property 7.87(d) gives

$$d(S,U) \leq \|S - T\| + \|T - U\| = d(S,T) + d(T,U).$$

Problem 7F.8. (a) Prove that if $T \in \mathcal{L}(V)$ and $\|I - T\| < 1$, then T is invertible.
(b) Suppose that $S \in \mathcal{L}(V)$ is invertible. Prove that if $T \in \mathcal{L}(V)$ and $\|S - T\| < 1/\|S^{-1}\|$, then T is invertible.
This exercise shows that the set of invertible operators in $\mathcal{L}(V)$ is an open subset of $\mathcal{L}(V)$, using the metric defined in Exercise 7.

Solution. (a) It suffices to show T is injective, since an injective operator on the finite-dimensional space V is invertible (3.65). If $Tv = 0$, then by 7.89,

$$\|v\| = \|v - Tv\| = \|(I - T)v\| \leq \|I - T\| \, \|v\|.$$

Were $v \neq 0$, dividing by $\|v\| > 0$ would give $1 \leq \|I - T\| < 1$. Hence $v = 0$.
(b) Since $V \neq \{0\}$ and $S^{-1} \neq 0$, we have $\|S^{-1}\| > 0$ by 7.87(b), so the hypothesis makes sense. From $I - S^{-1}T = S^{-1}(S - T)$ and submultiplicativity (Exercise 5 of this section),

$$\|I - S^{-1}T\| \leq \|S^{-1}\| \, \|S - T\| < \|S^{-1}\| \cdot \|S^{-1}\|^{-1} = 1.$$

So $S^{-1}T$ is invertible by part (a), and $T = S(S^{-1}T)$ is a composition of invertible operators, hence invertible.

Problem 7F.9. Suppose $T \in \mathcal{L}(V)$. Prove that for every $\epsilon > 0$, there exists an invertible operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \epsilon$.

Solution. Take $S = T - \lambda I$, where $\lambda \in (0, \epsilon)$ is not an eigenvalue of T; such a λ exists because T has at most $\dim V$ distinct eigenvalues by 5.12, while $(0, \epsilon) \subseteq \mathbf{R} \subseteq \mathbf{F}$ is infinite.
Then S is injective, hence invertible by 3.65. Since $T - S = \lambda I$ and $\|I\| = 1$ by 7.90, property 7.87(c) gives

$$\|T - S\| = |\lambda| \, \|I\| = \lambda \in (0, \epsilon).$$

Problem 7F.10. Suppose $\dim V > 1$ and $T \in \mathcal{L}(V)$ is not invertible. Prove that for every $\epsilon > 0$, there exists $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \epsilon$ and S is not invertible.

Solution. Take $Sv = Tv + \delta\langle v, w\rangle w$ with $\delta = \epsilon/2$, where u, w are chosen as follows. Since T is not invertible it is not injective (3.65), so there is $u \in \operatorname{null} T$ with $\|u\| = 1$; and $\dim(\operatorname{span}(u))^{\perp} = \dim V - 1 \geq 1$ by 6.51, so there is $w \in V$ with $\|w\| = 1$ and $\langle u, w\rangle = 0$.
Then $Su = Tu + \delta\langle u, w\rangle w = 0$ with $u \neq 0$, so S is not injective, hence not invertible. Since $(S - T)v = \delta\langle v, w\rangle w$, the Cauchy-Schwarz inequality 6.14 gives

$$\|(S - T)v\| = \delta|\langle v, w\rangle| \leq \delta\|v\|\|w\| = \delta\|v\|,$$

so $\|S - T\| \leq \delta$ by 7.88(c); and $(S - T)w = \delta w$ with $\|w\| = 1$ gives $\|S - T\| \geq \delta$. Hence $\|T - S\| = \|S - T\| = \delta$ by 7.87(c), so $0 < \|T - S\| < \epsilon$.

Problem 7F.11. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that for every $\epsilon > 0$ there exists a diagonalizable operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \epsilon$.

Solution. Take S to be T with its diagonal entries nudged to be distinct. Precisely, let $n = \dim V$; since $\mathbf{F} = \mathbf{C}$, Schur's theorem 6.38 gives an orthonormal basis $e_1, \ldots, e_n$ of V for which $A = \mathcal{M}(T, (e_1, \ldots, e_n))$ is upper triangular, with diagonal entries $\lambda_1, \ldots, \lambda_n$.

Choose $\delta_1, \dots, \delta_n$ in the infinite set $D = \{\delta \in \mathbf{C} : 0 < |\delta| < \epsilon\}$ inductively so that the numbers $\lambda_k + \delta_k$ are distinct: having chosen $\delta_1, \dots, \delta_{k-1}$, each $j < k$ forbids the single value $\delta = \lambda_j + \delta_j - \lambda_k$, so at most $k-1$ elements of D are excluded. Let S be the operator whose matrix with respect to $e_1, \dots, e_n$ is A with λ_k replaced by $\lambda_k + \delta_k$ for each k.
That matrix is still upper triangular, so the eigenvalues of S are exactly $\lambda_1 + \delta_1, \dots, \lambda_n + \delta_n$ by 5.41; these are $n = \dim V$ distinct numbers, so S is diagonalizable by 5.58. Only the diagonal entries changed, so $(T - S)e_k = -\delta_k e_k$ for each k. Picking m with $|\delta_m| = \max_k |\delta_k|$ and writing $v = a_1 e_1 + \cdots + a_n e_n$, orthonormality gives

$$\|(T - S)v\|^2 = \sum_{k=1}^{n} |\delta_k|^2 |a_k|^2 \le |\delta_m|^2 \|v\|^2,$$

so $\|T - S\| \le |\delta_m|$ by 7.88(c), while $\|(T - S)e_m\| = |\delta_m|$ with $\|e_m\| = 1$ gives the reverse. Hence $\|T - S\| = |\delta_m| \in (0, \epsilon)$.

Problem 7F.12. Suppose $T \in \mathcal{L}(V)$ is a positive operator. Show that $\|\sqrt{T}\| = \sqrt{\|T\|}$.

Solution. By 7.38 ((a) $\Rightarrow$ (c)) there is an orthonormal basis $e_1, \dots, e_n$ of V with $Te_k = \lambda_k e_k$ and each $\lambda_k \ge 0$. The operator R determined by $Re_k = \sqrt{\lambda_k}\, e_k$ has a diagonal matrix with nonnegative entries with respect to this orthonormal basis, hence is positive by 7.38 ((c) $\Rightarrow$ (a)), and $R^2 e_k = \lambda_k e_k = Te_k$ for each k, so $R^2 = T$. By uniqueness of the positive square root (7.39), $\sqrt{T} = R$.
An operator whose matrix with respect to an orthonormal basis is diagonal with entries $\mu_1, \dots, \mu_n$ has norm $\max_k |\mu_k|$ (third bullet of Example 7.90). Applying this to T and to R, and using that $t \mapsto \sqrt{t}$ is increasing on $[0, \infty)$,

$$\|\sqrt{T}\| = \max_k \sqrt{\lambda_k} = \sqrt{\max_k \lambda_k} = \sqrt{\|T\|}.$$

Method (2): Exercise 19 of this section applied to $A = \sqrt{T}$ reads $\|A^* A\| = \|A\|^2$; since $\sqrt{T}$ is self-adjoint with $(\sqrt{T})^2 = T$, this is $\|T\| = \|\sqrt{T}\|^2$.

Problem 7F.13. Suppose $S, T \in \mathcal{L}(V)$ are positive operators. Show that

$$\|S - T\| \le \max\{\|S\|, \|T\|\} \le \|S + T\|.$$

Solution. Lemma. If $A \in \mathcal{L}(V)$ is self-adjoint, then $|\langle Au, u\rangle| \le \|A\|$ for every unit vector u, with equality for some unit vector.
Indeed, for $\|u\| = 1$ the Cauchy-Schwarz inequality 6.14 and 7.89 give $|\langle Au, u\rangle| \le \|Au\|\,\|u\| \le \|A\|$. For equality, the spectral theorem (7.29 if $\mathbf{F} = \mathbf{R}$; 7.31 if $\mathbf{F} = \mathbf{C}$, a self-adjoint operator being normal) gives an orthonormal basis $e_1, \dots, e_n$ with $Ae_k = \mu_k e_k$, the μ_k real by 7.12; then $\|A\| = \max_k |\mu_k| = |\mu_m|$ for some m by the third bullet of Example 7.90, and $|\langle Ae_m, e_m\rangle| = |\mu_m| = \|A\|$.
Consequently, for positive R and unit u we have $0 \le \langle Ru, u\rangle \le \|R\|$, the left inequality by the definition 7.34 and the right by the lemma (such an $\langle Ru, u\rangle$ is a nonnegative real, so it equals its absolute value).
(i) $\|S - T\| \le \max\{\|S\|, \|T\|\}$. Since $(S - T)^* = S - T$, the lemma gives a unit vector v with $\|S - T\| = |a - b|$, where $a = \langle Sv, v\rangle$ and $b = \langle Tv, v\rangle$ are reals with $0 \le a \le \|S\|$ and $0 \le b \le \|T\|$. Then $a - b \le a \le \|S\|$ and $b - a \le b \le \|T\|$, so $|a - b| \le \max\{\|S\|, \|T\|\}$.
(ii) $\max\{\|S\|, \|T\|\} \le \|S+T\|$. Both sides are symmetric in S and T, so assume $\max\{\|S\|, \|T\|\} = \|S\|$. The lemma gives a unit v with $|\langle Sv, v\rangle| = \|S\|$, and positivity of S upgrades this to $\langle Sv, v\rangle = \|S\|$. Hence, by 7.86 (valid as $\|v\| \le 1$), then 6.14, then positivity of T,

$$\begin{aligned} \|S + T\| &\ge \|(S + T)v\| \ge |\langle (S + T)v, v\rangle| \\ &= \langle Sv, v\rangle + \langle Tv, v\rangle \ge \langle Sv, v\rangle = \|S\|, \end{aligned}$$

the middle equality because $\langle Sv, v\rangle$ and $\langle Tv, v\rangle$ are nonnegative reals.

Problem 7F.14. Suppose U and W are subspaces of V such that $\|P_U - P_W\| < 1$. Prove that $\dim U = \dim W$.

Solution. The restriction $P_U|_W\colon W \to U$ is injective, and by symmetry so is $P_W|_U\colon U \to W$; the two inequalities $\dim W \le \dim U$ and $\dim U \le \dim W$ follow.
For the injectivity: P_U maps into U because $\operatorname{range} P_U = U$ by 6.57(d). If $w \in W$ and $P_U w = 0$, then $P_W w = w$ by 6.57(b), so by 7.89 and then 7.87(c),

$$\|w\| = \|(P_W - P_U)w\| \le \|P_W - P_U\|\,\|w\| = \|P_U - P_W\|\,\|w\|.$$

Were $w \neq 0$, dividing by $\|w\| > 0$ would give $1 \le \|P_U - P_W\| < 1$. Hence $w = 0$.
Now the fundamental theorem of linear maps 3.21 gives

$$\dim W = 0 + \dim \operatorname{range}\bigl(P_U|_W\bigr) \le \dim U,$$

since $\operatorname{range}(P_U|_W)$ is a subspace of U. The hypothesis is symmetric in U and W (again by 7.87(c)), so the same argument with the roles swapped gives $\dim U \le \dim W$.

Problem 7F.15. Define $T \in \mathcal{L}(\mathbf{F}^3)$ by

$$T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2).$$

Find (explicitly) a unitary operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T = S\sqrt{T^*T}$.

Solution. Take $S(z_1, z_2, z_3) = (z_3, z_1, z_2)$.
With e_1, e_2, e_3 the standard (orthonormal) basis, $Te_1 = 2e_2$, $Te_2 = 3e_3$, $Te_3 = e_1$. The adjoint identity gives $\langle T^*Te_j, e_k\rangle = \langle Te_j, Te_k\rangle$, and $2e_2, 3e_3, e_1$ is an orthogonal list with squared norms $4, 9, 1$; since a vector is determined by its inner products with an orthonormal basis (6.30),

$$T^*T(z_1, z_2, z_3) = (4z_1, 9z_2, z_3).$$

The operator $R(z_1, z_2, z_3) = (2z_1, 3z_2, z_3)$ is diagonal with nonnegative entries $2, 3, 1$ with respect to e_1, e_2, e_3, hence positive by 7.38 ((c) $\Rightarrow$ (a)), and $R^2 = T^*T$; by uniqueness of the positive square root (7.39), $\sqrt{T^*T} = R$.
Now Se_1, Se_2, Se_3 is the list e_2, e_3, e_1, an orthonormal basis, so S is unitary by 7.53 ((d) $\Rightarrow$ (a)), and

$$S\sqrt{T^*T}(z_1, z_2, z_3) = S(2z_1, 3z_2, z_3) = (z_3, 2z_1, 3z_2) = T(z_1, z_2, z_3).$$

Problem 7F.16. Suppose $S \in \mathcal{L}(V)$ is a positive invertible operator. Prove that there exists $\delta > 0$ such that T is a positive operator for every self-adjoint operator $T \in \mathcal{L}(V)$ with $\|S - T\| < \delta$.

Solution. Take $\delta = \min\{\lambda_1, \ldots, \lambda_n\}$, where by 7.38 ((a) $\Rightarrow$ (c)) the λ_k are the eigenvalues of S along an orthonormal basis $e_1, \ldots, e_n$ of eigenvectors, $Se_k = \lambda_k e_k$ with $\lambda_k \ge 0$; invertibility of S rules out 0 as an eigenvalue, so $\delta > 0$.
Writing $v = a_1e_1 + \cdots + a_ne_n$ with $a_k = \langle v, e_k\rangle$ (6.30(a)) and using orthonormality and Parseval's identity 6.30(b),

$$\begin{aligned}\langle Sv, v\rangle &= \sum_{k=1}^{n} \lambda_k |a_k|^2 \\ &\ge \delta \sum_{k=1}^{n} |a_k|^2 = \delta\|v\|^2.\end{aligned}$$

Now suppose T is self-adjoint with $\|S - T\| < \delta$, and let $v \in V$. Then $S - T$ is self-adjoint, so

$\langle (S-T)v, v\rangle$ is real (7.14 when $\mathbf{F} = \mathbf{C}$; automatic when $\mathbf{F} = \mathbf{R}$), and

$$\begin{aligned}
\langle Tv, v\rangle &= \langle Sv, v\rangle - \langle (S-T)v, v\rangle \\
&\geq \delta \|v\|^2 - \big|\langle (S-T)v, v\rangle\big| \\
&\geq \delta \|v\|^2 - \|(S-T)v\|\,\|v\| \\
&\geq \delta \|v\|^2 - \|S-T\|\,\|v\|^2 \\
&= \big(\delta - \|S-T\|\big)\|v\|^2 \\
&\geq 0,
\end{aligned}$$

where the second inequality is the Cauchy–Schwarz inequality 6.14, the third uses 7.89, and the last holds because $\|S-T\| < \delta$. Thus T is self-adjoint with $\langle Tv, v\rangle \geq 0$ for all v, hence positive by 7.34.

Problem 7F.17. Prove that if $u \in V$ and φ_u is the linear functional on V defined by the equation $\varphi_u(v) = \langle v, u\rangle$, then $\|\varphi_u\| = \|u\|$.
Here we are thinking of the scalar field $\mathbf{F}$ as an inner product space with $\langle \alpha, \beta\rangle = \alpha\overline{\beta}$ for all $\alpha, \beta \in \mathbf{F}$. Thus $\|\varphi_u\|$ means the norm of φ_u as a linear map from V to $\mathbf{F}$.

Solution. The norm on $\mathbf{F}$ is the absolute value, $\|\alpha\| = \sqrt{\alpha\overline{\alpha}} = |\alpha|$, so by 7.86

$$\|\varphi_u\| = \max\{|\langle v, u\rangle| : v \in V,\ \|v\| \leq 1\}.$$

(i) $u = 0$: then $\varphi_u = 0$, so $\|\varphi_u\| = 0 = \|u\|$.
(ii) $u \neq 0$: for $\|v\| \leq 1$ the Cauchy–Schwarz inequality 6.14 gives $|\langle v, u\rangle| \leq \|v\|\,\|u\| \leq \|u\|$, so $\|\varphi_u\| \leq \|u\|$; and the unit vector $v = u/\|u\|$ has

$$|\varphi_u(v)| = \frac{|\langle u, u\rangle|}{\|u\|} = \frac{\|u\|^2}{\|u\|} = \|u\|,$$

so $\|\varphi_u\| \geq \|u\|$. Hence $\|\varphi_u\| = \|u\|$.

Problem 7F.18. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $T \in \mathcal{L}(V, W)$.
(a) Prove that $\max\{\|Te_1\|, \dots, \|Te_n\|\} \leq \|T\| \leq (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$.
(b) Prove that $\|T\| = (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$ if and only if $\dim \operatorname{range} T \leq 1$.
Here $e_1, \dots, e_n$ is an arbitrary orthonormal basis of V, not necessarily connected with a singular value decomposition of T. If $s_1, \dots, s_n$ is the list of singular values of T, then the right side of the inequality above equals $(s_1^2 + \cdots + s_n^2)^{1/2}$, as was shown in Exercise 11(a) in Section 7E.

Solution. Write $c = \big(\|Te_1\|^2 + \cdots + \|Te_n\|^2\big)^{1/2}$.
(a) Each e_k is a unit vector, so $\|Te_k\| \leq \|T\|$ by 7.88(b); taking the maximum over k gives the left inequality.
For the right inequality, let $\|v\| \leq 1$. By 6.30(a), $Tv = \sum_{k=1}^n \langle v, e_k\rangle\, Te_k$, so the triangle inequality 6.17 and then the Cauchy–Schwarz inequality in $\mathbf{R}^n$ applied to $\big(|\langle v, e_k\rangle|\big)_k$ and $\big(\|Te_k\|\big)_k$ give

$$\begin{aligned}
\|Tv\| &\leq \sum_{k=1}^n |\langle v, e_k\rangle|\,\|Te_k\| \\
&\leq \Big(\sum_{k=1}^n |\langle v, e_k\rangle|^2\Big)^{1/2}\Big(\sum_{k=1}^n \|Te_k\|^2\Big)^{1/2} = \|v\|\,c \leq c,
\end{aligned}$$

where the equality uses Parseval's identity 6.30(b). Hence $\|T\| \leq c$ by 7.86.
(b) Claim (equality in the triangle inequality). If $x_1, \dots, x_n \in W$ satisfy $\|x_1 + \cdots + x_n\| = \|x_1\| + \cdots + \|x_n\|$, then all the x_k lie in a common one-dimensional span.
Indeed, put $x = x_1 + \cdots + x_n$. If $x = 0$ then $\sum_k \|x_k\| = 0$, so every $x_k = 0$. Otherwise set $w = x/\|x\|$;

then

$$\sum_{k=1}^{n} \|x_k\| = \|x\| = \langle x, w \rangle = \sum_{k=1}^{n} \langle x_k, w \rangle,$$

and taking real parts, the termwise bounds $\operatorname{Re}\langle x_k, w\rangle \le |\langle x_k, w\rangle| \le \|x_k\|$ (6.14) must all be equalities. So $|\langle x_k, w\rangle| = \|x_k\|\,\|w\|$, which by the equality condition in 6.14 makes one of x_k, w a scalar multiple of the other; as $w \neq 0$, either way $x_k \in \operatorname{span}(w)$.
(i) Suppose $\dim \operatorname{range} T \le 1$. If $T = 0$ then $\|T\| = 0 = c$. Otherwise choose $w \in W$ with $\|w\| = 1$ and $\operatorname{range} T = \operatorname{span}(w)$, and write $Te_k = \alpha_k w$, so $\|Te_k\| = |\alpha_k|$ and $c = (\sum_k |\alpha_k|^2)^{1/2} > 0$. Define

$$v = \frac{1}{c} \sum_{k=1}^{n} \overline{\alpha_k}\, e_k.$$

By Parseval's identity, $\|v\| = \frac{1}{c}(\sum_k |\alpha_k|^2)^{1/2} = 1$. Moreover

$$Tv = \frac{1}{c} \sum_k \overline{\alpha_k}\, Te_k = \frac{1}{c}\Big(\sum_k |\alpha_k|^2\Big) w = c\, w,$$

so $\|Tv\| = c$. Hence $\|T\| \ge c$, which with (a) gives $\|T\| = c$.
(ii) Conversely, suppose $\|T\| = c$. If $c = 0$ then $T = 0$ by 7.87(b) and $\dim \operatorname{range} T = 0$. So assume $c > 0$. By 7.88(b) there is v with $\|v\| = 1$ and $\|Tv\| = c$. Set $a_k = |\langle v, e_k\rangle|$ and $b_k = \|Te_k\|$; then $\sum_k a_k^2 = 1$ by 6.30(b) and $\sum_k b_k^2 = c^2$, and the chain in part (a) applied to this v reads

$$\begin{aligned} c = \|Tv\| &\le \sum_{k=1}^{n} a_k b_k \\ &\le \Big(\sum_{k=1}^{n} a_k^2\Big)^{1/2} \Big(\sum_{k=1}^{n} b_k^2\Big)^{1/2} = 1 \cdot c = c, \end{aligned}$$

so both inequalities are equalities.
Equality in the second says $\langle a, b\rangle = \|a\|\,\|b\|$ for the nonzero vectors $a = (a_k)$, $b = (b_k)$ in $\mathbf{R}^n$, so by the equality condition in 6.14 we may write $a = \lambda b$ with $\lambda \neq 0$; then $1 = \|a\| = |\lambda| c$, and choosing j with $b_j > 0$ forces $\lambda = a_j/b_j > 0$. Hence $a = b/c$, that is,

$$|\langle v, e_k\rangle| = \frac{\|Te_k\|}{c} \quad \text{for } k = 1, \dots, n,$$

so $\langle v, e_k\rangle = 0$ if and only if $Te_k = 0$.
Equality in the first says $\big\|\sum_k x_k\big\| = \sum_k \|x_k\|$ for $x_k = \langle v, e_k\rangle\, Te_k$, so by the claim there is w with every $x_k \in \operatorname{span}(w)$. If $\langle v, e_k\rangle \neq 0$, dividing by it gives $Te_k \in \operatorname{span}(w)$; if $\langle v, e_k\rangle = 0$, then $Te_k = 0 \in \operatorname{span}(w)$. Since $Te_1, \dots, Te_n$ spans $\operatorname{range} T$, we get $\operatorname{range} T \subseteq \operatorname{span}(w)$ and $\dim \operatorname{range} T \le 1$.

Problem 7F.19. Prove that if $T \in \mathcal{L}(V, W)$, then $\|T^*T\| = \|T\|^2$.
This formula for $\|T^*T\|$ leads to the important subject of C^*-algebras.

Solution. ($\le$) Two applications of 7.89 give $\|T^*Tv\| \le \|T^*\|\,\|T\|\,\|v\|$ for all $v \in V$, and $\|T^*T\|$ is the smallest such constant by 7.88(c), so

$$\|T^*T\| \le \|T^*\|\,\|T\| = \|T\|^2,$$

the last equality because a linear map and its adjoint have the same norm (7.91).
($\ge$) If $\|v\| \le 1$, then the adjoint identity, the Cauchy–Schwarz inequality 6.14 and 7.89 give

$$\begin{aligned} \|Tv\|^2 = \langle v, T^*Tv\rangle &\le \|v\|\,\|T^*Tv\| \\ &\le \|T^*T\|\,\|v\|^2 \le \|T^*T\|. \end{aligned}$$

Taking the maximum over such v (7.86) gives $\|T\|^2 \le \|T^*T\|$.
Method (2): T^*T is positive and, by 7.65, its eigenvalues are the squares s_k^2 of the singular values of T; the singular values of a positive operator are its eigenvalues, so the largest singular value of T^*T is s_1^2 and 7.88(a) gives $\|T^*T\| = s_1^2 = \|T\|^2$.

Problem 7F.20. Suppose $T \in \mathcal{L}(V)$ is normal. Prove that $\|T^k\| = \|T\|^k$ for every positive integer k.

Solution. Submultiplicativity (Exercise 5 of this section) gives $\|A^k\| \le \|A\|^k$ for every $A \in \mathcal{L}(V)$, so only $\|T\|^k \le \|T^k\|$ needs proof. The argument below works over both $\mathbf{R}$ and $\mathbf{C}$ (over $\mathbf{R}$ a normal operator need not be diagonalizable, so the spectral theorem is unavailable).
Every power of T is normal: T commutes with T^*, hence T^j commutes with $(T^*)^j = (T^j)^*$ by repeated use of 7.5(d).
Next, $\|T^2\| = \|T\|^2$ for normal T. Using $(T^2)^* = (T^*)^2$ and $TT^* = T^*T$,

$$(T^2)^*T^2 = T^*(T^*T)T = T^*(TT^*)T = (T^*T)^2.$$

Put $A = T^*T$, a positive (hence self-adjoint) operator by 7.64(a), so $A^*A = A^2$. Exercise 19 of this section, applied in turn to T^2, to A, and to T, gives

$$\|T^2\|^2 = \left\|(T^2)^*T^2\right\| = \|A^*A\| = \|A\|^2 = \left(\|T\|^2\right)^2,$$

so $\|T^2\| = \|T\|^2$. Since each $R = T^{2^m}$ is normal, induction on m now gives $\|T^{2^m}\| = \|T\|^{2^m}$ for every $m \ge 0$ (the step being $\|R^2\| = \|R\|^2$).
For general k: if $T = 0$ both sides vanish, so assume $\|T\| > 0$ (7.87(b)) and pick $N = 2^m \ge k$. Then

$$\|T\|^N = \|T^N\| \le \|T^k\|\,\|T^{N-k}\| \le \|T^k\|\,\|T\|^{N-k},$$

which also holds when $N = k$ since $T^0 = I$ has norm 1. Dividing by $\|T\|^{N-k} > 0$ gives $\|T\|^k \le \|T^k\|$.

Problem 7F.21. Suppose $\dim V > 1$ and $\dim W > 1$. Prove that the norm on $\mathcal{L}(V, W)$ does not come from an inner product. In other words, prove that there does not exist an inner product on $\mathcal{L}(V, W)$ such that

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} = \sqrt{\langle T, T\rangle}$$

for all $T \in \mathcal{L}(V, W)$.

Solution. Take orthonormal lists e_1, e_2 in V and f_1, f_2 in W (available by Gram–Schmidt 6.32, since $\dim V \ge 2$ and $\dim W \ge 2$) and define $A, B \in \mathcal{L}(V, W)$ by

$$Av = \langle v, e_1\rangle f_1 \quad \text{and} \quad Bv = \langle v, e_2\rangle f_2;$$

these violate the parallelogram equality 6.21, which every inner-product norm obeys.
Indeed $\|Av\| = |\langle v, e_1\rangle| \le \|v\|$ by 6.14, so $\|A\| \le 1$ by 7.88(c), while $\|Ae_1\| = 1 = \|e_1\|$ gives $\|A\| \ge 1$; thus $\|A\| = 1$, and likewise $\|B\| = 1$. For $A \pm B$, orthonormality of f_1, f_2 and the Pythagorean theorem give

$$\|(A \pm B)v\|^2 = |\langle v, e_1\rangle|^2 + |\langle v, e_2\rangle|^2 \le \|v\|^2,$$

the inequality being Bessel's (extend e_1, e_2 to an orthonormal basis and apply Parseval's identity 6.30(b)); equality holds at $v = e_1$, so $\|A + B\| = \|A - B\| = 1$. Hence

$$\|A + B\|^2 + \|A - B\|^2 = 2 \ne 4 = 2\left(\|A\|^2 + \|B\|^2\right),$$

so no inner product on $\mathcal{L}(V, W)$ has the operator norm as its associated norm.

Problem 7F.22. Suppose $T \in \mathcal{L}(V, W)$. Let $n = \dim V$ and let $s_1 \geq \cdots \geq s_n$ denote the singular values of T. Prove that if $1 \leq k \leq n$, then

$$\min\{\|T|_U\| : U \text{ is a subspace of } V \text{ with } \dim U = k\} = s_{n-k+1}.$$

Solution. The minimum is attained at $U = \operatorname{span}(e_{n-k+1}, \ldots, e_n)$, where $e_1, \ldots, e_n$ is an orthonormal basis of V with $T^*Te_j = s_j^2 e_j$ for each j; such a basis exists by the spectral theorem applied to the positive operator T^*T (7.64(a)), and the s_j are the singular values of T by the definition 7.65 (this is step 7.72 in the proof of 7.70).
By 6.30(a), $T^*Tv = \sum_j s_j^2 \langle v, e_j\rangle e_j$ for $v \in V$, so

$$\|Tv\|^2 = \langle T^*Tv, v\rangle = \sum_j s_j^2 |\langle v, e_j\rangle|^2,$$

while $\|v\|^2 = \sum_j |\langle v, e_j\rangle|^2$ by Parseval's identity 6.30(b); in particular $\|Te_j\| = s_j$.
(i) This U has $\dim U = k$, and for $v \in U$ with $\|v\| \leq 1$ the terms with $j < n-k+1$ drop out, leaving (since $s_j \leq s_{n-k+1}$ for $j \geq n-k+1$)

$$\|Tv\|^2 = \sum_{j \geq n-k+1} s_j^2 |\langle v, e_j\rangle|^2 \leq s_{n-k+1}^2 \|v\|^2 \leq s_{n-k+1}^2.$$

So $\|T|_U\| \leq s_{n-k+1}$, with equality since $e_{n-k+1} \in U$ is a unit vector with $\|Te_{n-k+1}\| = s_{n-k+1}$.
(ii) Let U be any subspace with $\dim U = k$ and put $U' = \operatorname{span}(e_1, \ldots, e_{n-k+1})$. By 2.43, and since $\dim(U + U') \leq n$,

$$\dim(U \cap U') \geq k + (n-k+1) - n = 1,$$

so there is $v \in U \cap U'$ with $\|v\| = 1$. For it the terms with $j > n-k+1$ drop out, and $s_j \geq s_{n-k+1}$ for $j \leq n-k+1$, so

$$\|Tv\|^2 = \sum_{j \leq n-k+1} s_j^2 |\langle v, e_j\rangle|^2 \geq s_{n-k+1}^2 \|v\|^2 = s_{n-k+1}^2,$$

giving $\|T|_U\| \geq \|Tv\| \geq s_{n-k+1}$. Hence the minimum exists and equals s_{n-k+1}.

Problem 7F.23. Suppose $T \in \mathcal{L}(V, W)$. Show that T is uniformly continuous with respect to the metrics on V and W that arise from the norms on those spaces (see Exercise 23 in Section 6B).

Solution. Given $\varepsilon > 0$, take $\delta = \varepsilon/(\|T\| + 1) > 0$, where $\|T\|$ is defined by 7.86 because V is finite-dimensional (the $+1$ only avoids dividing by 0 when $T = 0$). If $u, v \in V$ satisfy $d_V(u, v) = \|u - v\| < \delta$, then by linearity of T and 7.89,

$$d_W(Tu, Tv) = \|T(u - v)\| \leq \|T\|\,\|u - v\| < \frac{\|T\|}{\|T\| + 1}\,\varepsilon \leq \varepsilon.$$

This δ depends only on ε and T, so T is uniformly continuous.

Problem 7F.24. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$\|T^{-1}\| = \|T\|^{-1} \iff \frac{T}{\|T\|} \text{ is a unitary operator.}$$

Solution. Set $S = T/\|T\|$, which makes sense because T is invertible, hence nonzero, so $\|T\| > 0$ by 7.87(b). By 7.87(c), $\|S\| = 1$, and $S^{-1} = \|T\|\,T^{-1}$ gives $\|S^{-1}\| = \|T\|\,\|T^{-1}\|$; so the assertion is that $\|S^{-1}\| = 1$ if and only if S is unitary.

($\Rightarrow$) Suppose $\|S^{-1}\| = 1$. For $v \in V$, applying 7.89 to S and then to S^{-1} with the vector Sv,

$$\|Sv\| \le \|S\|\,\|v\| = \|v\| = \|S^{-1}(Sv)\| \le \|S^{-1}\|\,\|Sv\| = \|Sv\|.$$

So $\|Sv\| = \|v\|$ for all v, making S an isometry (7.44) and, being invertible, unitary (7.51).
($\Leftarrow$) If S is unitary, then $S^{-1} = S^*$ is unitary by 7.53(c) and 7.53(f), hence an isometry, so $\|S^{-1}w\| = \|w\| = 1$ for every unit vector w and $\|S^{-1}\| = 1$ by 7.88(b). Thus $\|T\|\,\|T^{-1}\| = 1$.

Problem 7F.25. Fix $u, x \in V$ with $u \neq 0$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u\rangle x$ for every $v \in V$. Prove that

$$\sqrt{T^*T}\,v = \frac{\|x\|}{\|u\|}\langle v, u\rangle u$$

for every $v \in V$.

Solution. Let $Rv = \frac{\|x\|}{\|u\|}\langle v, u\rangle u$; since T^*T has exactly one positive square root (7.39, 7.40), it suffices to show R is positive with $R^2 = T^*T$.
First, for $v, w \in V$,

$$\langle Tv, w\rangle = \langle v, u\rangle\langle x, w\rangle = \big\langle v, \langle w, x\rangle u\big\rangle,$$

so $T^*w = \langle w, x\rangle u$ by uniqueness of the adjoint, and hence

$$T^*Tv = \langle v, u\rangle\, T^*x = \langle v, u\rangle\langle x, x\rangle u = \|x\|^2\langle v, u\rangle u.$$

Next, R is self-adjoint: as $\|x\|/\|u\|$ is real, both $\langle Rv, w\rangle$ and $\langle v, Rw\rangle$ equal $\frac{\|x\|}{\|u\|}\langle v, u\rangle\langle u, w\rangle$ (Check!). And R is positive, since

$$\langle Rv, v\rangle = \frac{\|x\|}{\|u\|}\langle v, u\rangle\overline{\langle v, u\rangle} = \frac{\|x\|}{\|u\|}\big|\langle v, u\rangle\big|^2 \ge 0.$$

Finally,

$$R(Rv) = \frac{\|x\|^2}{\|u\|^2}\langle v, u\rangle\langle u, u\rangle\, u = \|x\|^2\langle v, u\rangle u = T^*Tv,$$

so $R^2 = T^*T$ and therefore $R = \sqrt{T^*T}$.

Problem 7F.26. Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if there exists a unique unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

Solution. Write $R = \sqrt{T^*T}$, a well-defined positive operator since T^*T is positive [7.64(a)] with a unique positive square root (7.39, 7.40). The polar decomposition 7.93 supplies at least one unitary S with $T = SR$, so the content is that S is unique exactly when T is invertible.
Both directions use $\operatorname{null} R = \operatorname{null} T$, which follows from $\|Rv\| = \|Tv\|$: as R is self-adjoint with $R^2 = T^*T$,

$$\|Rv\|^2 = \langle R^2v, v\rangle = \langle T^*Tv, v\rangle = \|Tv\|^2.$$

($\Rightarrow$) If T is invertible, then $\operatorname{null} R = \operatorname{null} T = \{0\}$, so R is invertible by 3.65. Any unitary S_1, S_2 with $S_1R = T = S_2R$ then satisfy $S_1 = TR^{-1} = S_2$, so $S = T(\sqrt{T^*T})^{-1}$ is the unique such operator.
($\Leftarrow$) Contrapositive: suppose T is not invertible, so $U := \operatorname{null} T \neq \{0\}$ and $\operatorname{null} R = U$. Fix a unitary S with $T = SR$ (7.93). Since R is self-adjoint, 7.6(d) gives

$$\operatorname{range} R = (\operatorname{null} R^*)^\perp = U^\perp,$$

and $V = U \oplus U^\perp$ by 6.49, so $Q(v_1 + v_2) = -v_1 + v_2$ (for $v_1 \in U$, $v_2 \in U^\perp$) defines an operator on V. By the Pythagorean theorem 6.12, $\|Q(v_1 + v_2)\|^2 = \|v_1\|^2 + \|v_2\|^2 = \|v_1 + v_2\|^2$, so Q is an isometry (7.44), hence unitary (7.51, using 3.65).
Put $S' = SQ$, unitary as a composition of unitary operators. Every Rv lies in $\operatorname{range} R = U^\perp$, so $Q(Rv) = Rv$ and hence $S'Rv = S(Rv) = Tv$. But $S' \neq S$, since $SQ = S$ would force $Q = I$,

contradicting $Qv_1 = -v_1 \neq v_1$ for nonzero $v_1 \in U$. So the unitary operator is not unique.

Problem 7F.27. Suppose $T \in \mathcal{L}(V)$ and $s_1, \ldots, s_n$ are the singular values of T. Let $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ be orthonormal bases of V such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for all $v \in V$. Define $S \in \mathcal{L}(V)$ by

$$Sv = \langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n.$$

(a) Show that S is unitary and $\|T - S\| = \max\{|s_1 - 1|, \ldots, |s_n - 1|\}$.
(b) Show that if $E \in \mathcal{L}(V)$ is unitary, then $\|T - E\| \geq \|T - S\|$.
[This exercise finds a unitary operator S that is as close as possible (among the unitary operators) to a given operator T.]

Solution. Write $M = \max\{|s_1 - 1|, \ldots, |s_n - 1|\}$ and fix k with $|s_k - 1| = M$.
(a) S is an isometry: for $v \in V$, orthonormality of $f_1, \ldots, f_n$ and then Parseval's identity 6.30(b) for $e_1, \ldots, e_n$ give

$$\|Sv\|^2 = \sum_{j=1}^{n} \big|\langle v, e_j\rangle\big|^2 = \|v\|^2.$$

Hence S is injective, so invertible (3.65), so unitary (7.51).
Subtracting the two displayed formulas gives

$$(T - S)v = \sum_{j=1}^{n} (s_j - 1)\langle v, e_j\rangle f_j,$$

so the same two orthonormality facts yield

$$\begin{aligned}\|(T - S)v\|^2 &= \sum_{j=1}^{n} (s_j - 1)^2 \big|\langle v, e_j\rangle\big|^2 \\ &\leq M^2 \sum_{j=1}^{n} \big|\langle v, e_j\rangle\big|^2 = M^2\|v\|^2,\end{aligned}$$

whence $\|T - S\| \leq M$. Conversely $(T - S)e_k = (s_k - 1)f_k$ has norm M while $\|e_k\| = 1$, so $\|T - S\| \geq M$ by the definition 7.86. Thus $\|T - S\| = M$.
(b) Let E be unitary. Then $\|Ee_k\| = \|e_k\| = 1$, and $Te_k = s_k f_k$ gives $\|Te_k\| = s_k$. The reverse triangle inequality (from 6.17) applied to Te_k and Ee_k gives

$$\|(T - E)e_k\| \geq \big|\, \|Te_k\| - \|Ee_k\| \,\big| = |s_k - 1| = M.$$

Since $\|e_k\| = 1$, by 7.86 and part (a),

$$\|T - E\| \geq \|(T - E)e_k\| \geq M = \|T - S\|.$$

Problem 7F.28. Suppose $T \in \mathcal{L}(V)$. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = \sqrt{TT^*}\, S$.

Solution. Take $S = R^*$, where R is the unitary operator given by the polar decomposition 7.93 applied to T^*:

$$T^* = R\sqrt{(T^*)^*T^*} = R\sqrt{TT^*},$$

using $(T^*)^* = T$ by 7.5(c). Since TT^* is positive by 7.64(a), so is $\sqrt{TT^*}$; in particular $\sqrt{TT^*}$ is self-adjoint. Taking adjoints and using 7.5(d),

$$T = \left(R\sqrt{TT^*}\right)^* = \left(\sqrt{TT^*}\right)^* R^* = \sqrt{TT^*}\, S.$$

Finally $S = R^*$ is unitary because R is, by the equivalence of (a) and (f) in 7.53.

Problem 7F.29. Suppose $T \in \mathcal{L}(V)$.
(a) Use the polar decomposition to show that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $TT^* = ST^*TS^*$.
(b) Show how (a) implies that T and T^* have the same singular values.

Solution. (a) Take the unitary S of the polar decomposition 7.93, so $T = SR$ with $R = \sqrt{T^*T}$. Then $R^* = R$ (positive operators are self-adjoint) and $R^2 = T^*T$, so $T^* = RS^*$ and

$$TT^* = (SR)(RS^*) = SR^2S^* = S(T^*T)S^*.$$

(b) By 7.65 the singular values of T are the nonnegative square roots of the eigenvalues of T^*T repeated by eigenspace dimension, and those of T^* are the same for $(T^*)^*T^* = TT^*$; so it suffices that T^*T and TT^* have equal eigenspace dimensions at every λ. Since $S^* = S^{-1}$ (7.53), part (a) reads $TT^* = S(T^*T)S^{-1}$, and for $v \in E(\lambda, T^*T)$,

$$(TT^*)(Sv) = S(T^*T)S^{-1}Sv = S(\lambda v) = \lambda(Sv),$$

so S maps $E(\lambda, T^*T)$ into $E(\lambda, TT^*)$. The same computation applied to $T^*T = S^{-1}(TT^*)S$ shows S^{-1} maps $E(\lambda, TT^*)$ into $E(\lambda, T^*T)$, so these restrictions are mutually inverse bijections and

$$\dim E(\lambda, T^*T) = \dim E(\lambda, TT^*) \quad \text{for all } \lambda \in \mathbf{F}.$$

Taking nonnegative square roots and listing in decreasing order gives the same list for T and for T^*.

Problem 7F.30. Suppose $T \in \mathcal{L}(V)$, $S \in \mathcal{L}(V)$ is a unitary operator, and $R \in \mathcal{L}(V)$ is a positive operator such that $T = SR$. Prove that $R = \sqrt{T^*T}$.
[This exercise shows that if we write T as the product of a unitary operator and a positive operator (as in the polar decomposition 7.93), then the positive operator equals $\sqrt{T^*T}$.]

Solution. R is a positive square root of T^*T, and T^*T has only one such (7.39, its positivity by 7.64(a)). Indeed R is positive by hypothesis, hence self-adjoint; and $S^*S = I$ by 7.53, so taking adjoints in $T = SR$ gives $T^* = RS^*$ and

$$T^*T = (RS^*)(SR) = R(S^*S)R = R^2.$$

Uniqueness in 7.39 now gives $R = \sqrt{T^*T}$.

Problem 7F.31. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$ and such that S and $\sqrt{T^*T}$ both have diagonal matrices with respect to the same orthonormal basis of V.

Solution. Take $Se_k = \mu_k e_k$, where $e_1, \ldots, e_n$ is an orthonormal basis of eigenvectors of T with eigenvalues $\lambda_1, \ldots, \lambda_n$ (complex spectral theorem 7.31) and

$$\mu_k = \begin{cases} \lambda_k/|\lambda_k| & \text{if } \lambda_k \neq 0, \\ 1 & \text{if } \lambda_k = 0. \end{cases}$$

First, $\sqrt{T^*T}\, e_k = |\lambda_k| e_k$. Indeed $T^*e_k = \overline{\lambda_k}\, e_k$ by normality (7.21(e)), so $T^*Te_k = |\lambda_k|^2 e_k$; and the operator R with $Re_k = |\lambda_k| e_k$ is positive, since its matrix with respect to the orthonormal basis

$e_1, \dots, e_n$ is diagonal with nonnegative entries, whence R is self-adjoint and, for $v = \sum_k a_k e_k$,

$$\langle Rv, v\rangle = \sum_{k=1}^{n} |\lambda_k|\,|a_k|^2 \geq 0.$$

As $R^2 e_k = |\lambda_k|^2 e_k = T^*Te_k$, uniqueness of positive square roots (7.39) gives $R = \sqrt{T^*T}$.
Next, S is unitary: $|\mu_k| = 1$, so $Se_1, \dots, Se_n$ is again orthonormal (Check!), hence an orthonormal basis, and 7.53 applies. Finally, for each k,

$$S\sqrt{T^*T}\, e_k = \mu_k|\lambda_k|e_k = \lambda_k e_k = Te_k,$$

the middle equality holding in both cases of the definition of μ_k, so $T = S\sqrt{T^*T}$. Both S and $\sqrt{T^*T}$ are diagonal with respect to $e_1, \dots, e_n$, with diagonals $\mu_1, \dots, \mu_n$ and $|\lambda_1|, \dots, |\lambda_n|$.

Problem 7F.32. Suppose that $T \in \mathcal{L}(V, W)$ and $T \neq 0$. Let $s_1, \dots, s_m$ denote the positive singular values of T. Show that there exists an orthonormal basis $e_1, \dots, e_m$ of $(\operatorname{null} T)^\perp$ such that

$$T\left(E\left(\frac{e_1}{s_1}, \dots, \frac{e_m}{s_m}\right)\right)$$

equals the ball in $\operatorname{range} T$ of radius 1 centered at 0.

Solution. Take $e_1, \dots, e_m$ and $f_1, \dots, f_m$ from the singular value decomposition 7.70, so that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

for all $v \in V$; here $m \geq 1$ since $T \neq 0$ (7.68(b)). Orthonormality of $f_1, \dots, f_m$ and the Pythagorean theorem give the norm identity used throughout,

$$\|Tv\|^2 = \sum_{k=1}^{m} s_k^2 |\langle v, e_k\rangle|^2.$$

(i) $e_1, \dots, e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$. With $U = \operatorname{span}(e_1, \dots, e_m)$, each $s_k > 0$, so the norm identity gives

$$Tv = 0 \iff \langle v, e_k\rangle = 0 \text{ for all } k \iff v \in U^\perp;$$

thus $\operatorname{null} T = U^\perp$ and $(\operatorname{null} T)^\perp = (U^\perp)^\perp = U$ by 6.52.
(ii) $f_1, \dots, f_m$ is an orthonormal basis of $\operatorname{range} T$: the display gives $\operatorname{range} T \subseteq \operatorname{span}(f_1, \dots, f_m)$, while $Te_k = s_k f_k$ gives the reverse.
By the definition 7.96, applied in $(\operatorname{null} T)^\perp$ with the positive numbers $1/s_1, \dots, 1/s_m$, the ellipsoid is

$$E = \left\{v \in (\operatorname{null} T)^\perp : \sum\nolimits_k s_k^2 |\langle v, e_k\rangle|^2 < 1\right\}.$$

If $v \in E$ then $\|Tv\| < 1$ by the norm identity, so $T(E)$ lies in the ball. Conversely let $w \in \operatorname{range} T$ with $\|w\| < 1$ and set

$$v = \sum_{k=1}^{m} \frac{\langle w, f_k\rangle}{s_k} e_k \in (\operatorname{null} T)^\perp.$$

Then $\langle v, e_k\rangle = \langle w, f_k\rangle / s_k$, so by (ii) and Parseval's identity 6.30(b),

$$\sum_{k=1}^{m} s_k^2 |\langle v, e_k\rangle|^2 = \sum_{k=1}^{m} |\langle w, f_k\rangle|^2 = \|w\|^2 < 1,$$

giving $v \in E$; and $Tv = \sum_k \langle w, f_k\rangle f_k = w$. Hence $T(E)$ is exactly the ball in $\operatorname{range} T$ of radius 1 centered at 0.

8 Operators on Complex Vector Spaces

Contents

8.1 Exercises 8A

Problem 8A.1. Suppose $T \in \mathcal{L}(V)$. Prove that if $\dim \operatorname{null} T^4 = 8$ and $\dim \operatorname{null} T^6 = 9$, then $\dim \operatorname{null} T^m = 9$ for all integers $m \geq 5$.

Solution. $\dim \operatorname{null} T^5 = 9$. Indeed the chain $\operatorname{null} T^4 \subseteq \operatorname{null} T^5 \subseteq \operatorname{null} T^6$ of 8.1 forces

$$8 = \dim \operatorname{null} T^4 \leq \dim \operatorname{null} T^5 \leq \dim \operatorname{null} T^6 = 9,$$

and $\dim \operatorname{null} T^5 = 8$ would give $\operatorname{null} T^4 = \operatorname{null} T^5$ (equal dimensions, 2.39), hence $\operatorname{null} T^4 = \operatorname{null} T^6$ by 8.2 and $\dim \operatorname{null} T^6 = 8$, a contradiction.
Now $\operatorname{null} T^5 \subseteq \operatorname{null} T^6$ with both of dimension 9, so $\operatorname{null} T^5 = \operatorname{null} T^6$ by 2.39, and 8.2 gives $\operatorname{null} T^m = \operatorname{null} T^5$ for every $m \geq 5$. Taking dimensions, $\dim \operatorname{null} T^m = 9$.

Problem 8A.2. Suppose $T \in \mathcal{L}(V)$, m is a positive integer, $v \in V$, and $T^{m-1}v \neq 0$ but $T^m v = 0$. Prove that $v, Tv, T^2v, \ldots, T^{m-1}v$ is linearly independent.
[The result in this exercise is used in the proof of 8.45.]

Solution. Suppose $a_0 v + a_1 Tv + \cdots + a_{m-1}T^{m-1}v = 0$; we show $a_j = 0$ by induction on j. Note first that $T^k v = T^{k-m}(T^m v) = 0$ for every $k \geq m$.
Let $0 \leq j \leq m-1$ and assume $a_0 = \cdots = a_{j-1} = 0$ (vacuous for $j = 0$), so that

$$a_j T^j v + a_{j+1}T^{j+1}v + \cdots + a_{m-1}T^{m-1}v = 0.$$

Apply T^{m-1-j}: each term with index $k \geq j+1$ becomes $T^{m-1-j+k}v = 0$ since $m-1-j+k \geq m$, leaving $a_j T^{m-1}v = 0$. As $T^{m-1}v \neq 0$, we get $a_j = 0$, and the list is linearly independent.

Problem 8A.3. Suppose $T \in \mathcal{L}(V)$. Prove that

$$V = \operatorname{null} T \oplus \operatorname{range} T \iff \operatorname{null} T^2 = \operatorname{null} T.$$

Solution. Both directions turn on $(\operatorname{null} T) \cap (\operatorname{range} T) = \{0\}$, which by 1.46 is equivalent to the sum being direct.
($\Rightarrow$) Only $\operatorname{null} T^2 \subseteq \operatorname{null} T$ needs proof, the reverse being 8.1. If $T^2 v = 0$ then $Tv \in (\operatorname{null} T) \cap (\operatorname{range} T) = \{0\}$, so $v \in \operatorname{null} T$.
($\Leftarrow$) If $v = Tu$ and $Tv = 0$, then $T^2 u = 0$, so $u \in \operatorname{null} T^2 = \operatorname{null} T$ and $v = Tu = 0$; hence the intersection is $\{0\}$ and the sum is direct. By 3.94 and the fundamental theorem of linear maps 3.21,

$$\dim(\operatorname{null} T \oplus \operatorname{range} T) = \dim \operatorname{null} T + \dim \operatorname{range} T = \dim V,$$

so this subspace is all of V by 2.39.

Problem 8A.4. Suppose $T \in \mathcal{L}(V)$, $\lambda \in \mathbf{F}$, and m is a positive integer such that the minimal polynomial of T is a polynomial multiple of $(z - \lambda)^m$. Prove that

$$\dim \operatorname{null}(T - \lambda I)^m \geq m.$$

Solution. Writing $S = T - \lambda I$, each of the m inclusions in the chain

$$\{0\} = \operatorname{null} S^0 \subseteq \cdots \subseteq \operatorname{null} S^m$$

of 8.1 is strict, whence $\dim \operatorname{null} S^m \geq m$.
Suppose instead $\operatorname{null} S^j = \operatorname{null} S^{j+1}$ for some $0 \leq j \leq m-1$, so $\operatorname{null} S^m = \operatorname{null} S^j$ by 8.2. Write $p(z) = (z - \lambda)^m q(z)$ for the minimal polynomial p of T; p is monic (5.22), so q is monic too. Put

$r(z) = (z - \lambda)^j q(z)$. For $v \in V$, set $u = q(T)v$; then by 5.17(a),

$$S^m u = (T - \lambda I)^m q(T)v = p(T)v = 0,$$

so $u \in \operatorname{null} S^m = \operatorname{null} S^j$ and therefore $r(T)v = S^j u = 0$. Thus r is monic with $r(T) = 0$ and

$$\deg r = j + \deg q < m + \deg q = \deg p,$$

contradicting the minimality in 5.22.

Problem 8A.5. Suppose $T \in \mathcal{L}(V)$ and m is a positive integer. Prove that

$$\dim \operatorname{null} T^m \leq m \dim \operatorname{null} T.$$

[Hint: Exercise 21 in Section 3B may be useful.]

Solution. It suffices to prove the one-step estimate

$$\dim \operatorname{null} T^k \leq \dim \operatorname{null} T^{k-1} + \dim \operatorname{null} T \qquad (k \geq 1),$$

since iterating it from $\dim \operatorname{null} T^1 = \dim \operatorname{null} T$ gives the claim by induction on m.
Let $\varphi\colon \operatorname{null} T^k \to \operatorname{null} T$ be $\varphi(v) = T^{k-1}v$, a linear map into $\operatorname{null} T$ because $T(T^{k-1}v) = T^k v = 0$. Its null space is

$$(\operatorname{null} T^k) \cap (\operatorname{null} T^{k-1}) = \operatorname{null} T^{k-1}$$

by 8.1, and $\dim \operatorname{range} \varphi \leq \dim \operatorname{null} T$ by 2.37. So the fundamental theorem of linear maps 3.21 applied to φ gives

$$\begin{aligned} \dim \operatorname{null} T^k &= \dim \operatorname{null} T^{k-1} + \dim \operatorname{range} \varphi \\ &\leq \dim \operatorname{null} T^{k-1} + \dim \operatorname{null} T. \end{aligned}$$

Problem 8A.6. Suppose $T \in \mathcal{L}(V)$. Show that

$$V = \operatorname{range} T^0 \supseteq \operatorname{range} T^1 \supseteq \cdots \supseteq \operatorname{range} T^k \supseteq \operatorname{range} T^{k+1} \supseteq \cdots.$$

Solution. $\operatorname{range} T^0 = \operatorname{range} I = V$, and for every nonnegative integer k and every $w = T^{k+1}v \in \operatorname{range} T^{k+1}$,

$$w = T^{k+1}v = T^k(Tv) \in \operatorname{range} T^k,$$

so $\operatorname{range} T^{k+1} \subseteq \operatorname{range} T^k$. Taking $k = 0, 1, 2, \ldots$ gives the chain.

Problem 8A.7. Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer such that

$$\operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Prove that $\operatorname{range} T^k = \operatorname{range} T^m$ for all $k > m$.

Solution. It suffices to show $\operatorname{range} T^{m+j} = \operatorname{range} T^{m+j+1}$ for every $j \geq 0$, since induction on j then gives $\operatorname{range} T^k = \operatorname{range} T^m$ for all $k \geq m$.
The inclusion $\operatorname{range} T^{m+j+1} \subseteq \operatorname{range} T^{m+j}$ is Exercise 6 in this section. Conversely let $w = T^{m+j}u = T^j(T^m u)$. By hypothesis $T^m u \in \operatorname{range} T^m = \operatorname{range} T^{m+1}$, so $T^m u = T^{m+1}x$ for some $x \in V$, and

$$w = T^j\big(T^{m+1}x\big) = T^{m+j+1}x \in \operatorname{range} T^{m+j+1}.$$

Problem 8A.8. Suppose $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{range} T^{\dim V} = \operatorname{range} T^{\dim V+1} = \operatorname{range} T^{\dim V+2} = \cdots .$$

Solution. Put $n = \dim V$; we show $\operatorname{range} T^k = \operatorname{range} T^n$ for every $k \geq n$. The inclusion $\operatorname{range} T^k \subseteq \operatorname{range} T^n$ is Exercise 6 in this section. By 8.3, $\dim \operatorname{null} T^k = \dim \operatorname{null} T^n$, so the fundamental theorem of linear maps 3.21 applied to T^k and to T^n gives

$$\dim \operatorname{range} T^k = n - \dim \operatorname{null} T^k = n - \dim \operatorname{null} T^n = \dim \operatorname{range} T^n.$$

A subspace of $\operatorname{range} T^n$ of that dimension is all of it (2.39).

Problem 8A.9. Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer. Prove that

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} \iff \operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Solution. Both equalities are equivalent to the corresponding equality of dimensions, because $\operatorname{null} T^m \subseteq \operatorname{null} T^{m+1}$ (8.1) and $\operatorname{range} T^{m+1} \subseteq \operatorname{range} T^m$ (Exercise 6 in this section), and a subspace of the same dimension is the whole space (2.39).
Subtracting the two instances of the fundamental theorem of linear maps 3.21,

$$\begin{aligned} \dim \operatorname{null} T^m + \dim \operatorname{range} T^m &= \dim V, \\ \dim \operatorname{null} T^{m+1} + \dim \operatorname{range} T^{m+1} &= \dim V, \end{aligned}$$

gives

$$\dim \operatorname{null} T^{m+1} - \dim \operatorname{null} T^m = \dim \operatorname{range} T^m - \dim \operatorname{range} T^{m+1},$$

so one side vanishes exactly when the other does.

Problem 8A.10. Define $T \in \mathcal{L}(\mathbf{C}^2)$ by $T(w,z) = (z,0)$. Find all generalized eigenvectors of T.

Solution. Every nonzero vector of $\mathbf{C}^2$ is a generalized eigenvector of T, each corresponding to the eigenvalue 0.
Indeed $T^2(w,z) = T(z,0) = (0,0)$, so T is nilpotent and 0 is its only eigenvalue (8.17(a)). Hence a nonzero v is a generalized eigenvector exactly when $(T - 0I)^2 v = T^2 v = 0$, which holds for every v.

Problem 8A.11. Suppose that $T \in \mathcal{L}(V)$. Prove that there is a basis of V consisting of generalized eigenvectors of T if and only if the minimal polynomial of T equals $(z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.
Assume $\mathbf{F} = \mathbf{R}$ because the case $\mathbf{F} = \mathbf{C}$ follows from 5.27(b) and 8.9.
This exercise states that the condition for there to be a basis of V consisting of generalized eigenvectors of T is the same as the condition for there to be a basis with respect to which T has an upper-triangular matrix (see 5.44).
Caution: If T has an upper-triangular matrix with respect to a basis $v_1, \ldots, v_n$ of V, then v_1 is an eigenvector of T but it is not necessarily true that $v_2, \ldots, v_n$ are generalized eigenvectors of T.

Solution. Write $n = \dim V \geq 1$, say that a monic $p \in \mathcal{P}(\mathbf{F})$ *splits* if $p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$ with all $\lambda_j \in \mathbf{F}$, and recall from the remark after 8.8 that a nonzero v is a generalized eigenvector for the eigenvalue λ exactly when $(T - \lambda I)^n v = 0$.
Lemma: a monic nonconstant $q \in \mathcal{P}(\mathbf{F})$ dividing a split monic p splits. Indeed, by 4.13 over $\mathbf{C}$ we may write $q(z) = (z - \alpha_1) \cdots (z - \alpha_r)$ with $\alpha_j \in \mathbf{C}$; each α_j is a zero of q, hence of p, and the zeros of a split p are its $\lambda_i \in \mathbf{F}$.
($\Leftarrow$) Induct on n. For $n = 1$ any nonzero vector is an eigenvector, hence a basis of generalized

eigenvectors. Let $n > 1$ and let $p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$ be the minimal polynomial of T, which is nonconstant, so $m \geq 1$; then λ_1 is an eigenvalue of T by 5.27(a). By 8.4 applied to $T - \lambda_1 I$,

$$V = \operatorname{null}(T - \lambda_1 I)^n \oplus \operatorname{range}(T - \lambda_1 I)^n.$$

(i) If $\operatorname{null}(T - \lambda_1 I)^n = V$, every basis of V consists of generalized eigenvectors for λ_1.
(ii) Otherwise put $U = \operatorname{range}(T - \lambda_1 I)^n \neq \{0\}$; since λ_1 is an eigenvalue, the null space is also nonzero, so $0 < \dim U < n$. By 5.18 (with the polynomial $(z - \lambda_1)^n$) U is invariant under T; set $S = T|_U$. By 5.31 the minimal polynomial of S divides p, hence splits by the lemma, so induction gives a basis $u_1, \ldots, u_k$ of U of generalized eigenvectors of S. Each u_j is one for T as well, since $(S - \mu I_U)^r u_j = 0$ gives $(T - \mu I)^r u_j = 0$ and an eigenvector of S is one of T. Adjoining a basis $w_1, \ldots, w_l$ of $\operatorname{null}(T - \lambda_1 I)^n$, whose vectors are generalized eigenvectors for λ_1, gives the required basis of V by the direct sum above.
($\Rightarrow$) Let $v_1, \ldots, v_n$ be a basis of generalized eigenvectors, let $\mu_1, \ldots, \mu_k \in \mathbf{F}$ be the distinct eigenvalues of T (at least one exists), and set $p(z) = (z - \mu_1)^n \cdots (z - \mu_k)^n$, which is monic and splits. Each v_j corresponds to exactly one μ_i (8.11), so $(T - \mu_i I)^n v_j = 0$; as the factors of $p(T)$ commute (5.17), $p(T)v_j = 0$ for every j and hence $p(T) = 0$. By 5.29 the minimal polynomial of T divides p, so it splits by the lemma.

Problem 8A.12. Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector in V is a generalized eigenvector of T. Prove that there exists $\lambda \in \mathbf{F}$ such that $T - \lambda I$ is nilpotent.

Solution. Take λ to be the unique eigenvalue of T; then $(T - \lambda I)^n = 0$ where $n = \dim V$, since every nonzero v is a generalized eigenvector and so satisfies $(T - \lambda I)^n v = 0$ (remark after 8.8), λ being the only eigenvalue available.
At least one eigenvalue exists because $V \neq \{0\}$ and a nonzero vector is by hypothesis a generalized eigenvector. For uniqueness, suppose $\alpha \neq \beta$ are eigenvalues, with eigenvectors u, w; these are linearly independent by 5.11, so $u + w \neq 0$ and $(T - \gamma I)^n (u + w) = 0$ for some eigenvalue γ. Since $(T - \gamma I)^n u = (\alpha - \gamma)^n u$ and likewise for w,

$$(\alpha - \gamma)^n u + (\beta - \gamma)^n w = 0,$$

so independence forces $\alpha = \gamma = \beta$, a contradiction.

Problem 8A.13. Suppose $S, T \in \mathcal{L}(V)$ and ST is nilpotent. Prove that TS is nilpotent.

Solution. Pick k with $(ST)^k = 0$. Regrouping the $2k + 2$ factors of $(TS)^{k+1}$ as one T, then k blocks ST, then one S,

$$(TS)^{k+1} = T(ST)^k S = 0,$$

so TS is nilpotent.

Problem 8A.14. Suppose $T \in \mathcal{L}(V)$ is nilpotent and $T \neq 0$. Prove T is not diagonalizable.

Solution. If T were diagonalizable, 5.55 would give a basis $v_1, \ldots, v_n$ of eigenvectors, $Tv_j = \lambda_j v_j$. Nilpotence forces $\lambda_j = 0$ for every j (8.17(a)), so T kills a basis of V and hence $T = 0$, contradicting $T \neq 0$.
Method (2): by 8.18 the minimal polynomial of T is z^m, while diagonalizability would make it a product of distinct factors $z - \lambda_i$ (5.62); comparing factorizations (4.13) gives $m = 1$, i.e. $T = 0$.

Problem 8A.15. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if every generalized eigenvector of T is an eigenvector of T.
For $\mathbf{F} = \mathbf{C}$, this exercise adds another equivalence to the list of conditions for diagonalizability in 5.55.

Solution. Write $n = \dim V$ and let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T.
($\Rightarrow$) Let v be a generalized eigenvector, corresponding to exactly one eigenvalue λ_j by 8.11, so $v \neq 0$ and $(T - \lambda_j I)^n v = 0$ (remark after 8.8). Diagonalizability gives $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$ by 5.55, so write $v = u_1 + \cdots + u_m$ with $u_k \in E(\lambda_k, T)$; then $(T - \lambda_j I)^n u_k = (\lambda_k - \lambda_j)^n u_k$ and

$$0 = (T - \lambda_j I)^n v = \sum_{k=1}^{m} (\lambda_k - \lambda_j)^n u_k.$$

The k-th summand lies in $E(\lambda_k, T)$, so directness (1.45) makes each vanish; as $(\lambda_k - \lambda_j)^n \neq 0$ for $k \neq j$, we get $v = u_j \in E(\lambda_j, T)$, an eigenvector.
($\Leftarrow$) Since $\mathbf{F} = \mathbf{C}$, by 8.9 there is a basis of V of generalized eigenvectors of T; by hypothesis it is a basis of eigenvectors, so T is diagonalizable by 5.55.

Problem 8A.16. (a) Give an example of nilpotent operators S, T on the same vector space such that neither $S + T$ nor ST is nilpotent.
(b) Suppose $S, T \in \mathcal{L}(V)$ are nilpotent and $ST = TS$. Prove that $S + T$ and ST are nilpotent.

Solution. (a) Take $S, T \in \mathcal{L}(\mathbf{F}^2)$ given by $S(x, y) = (y, 0)$ and $T(x, y) = (0, x)$, so $S^2 = T^2 = 0$. Then $(S + T)(x, y) = (y, x)$, so $(S + T)^2 = I$ and no power of $S + T$ is 0; and $(ST)(x, y) = (x, 0)$, so $(ST)^k = ST \neq 0$ for every $k \geq 1$.
(b) Pick m, n with $S^m = 0$ and $T^n = 0$. Commutativity lets all copies of S be moved left, so

$$(ST)^m = S^m T^m = 0.$$

Commutativity also makes the binomial theorem valid in $\mathcal{L}(V)$ for S and T:

$$(S + T)^{m+n-1} = \sum_{j=0}^{m+n-1} \binom{m+n-1}{j} S^j T^{m+n-1-j}.$$

Each term vanishes: if $j \geq m$ then $S^j = 0$, and if $j \leq m - 1$ then $m + n - 1 - j \geq n$, so $T^{m+n-1-j} = 0$. Hence $(S + T)^{m+n-1} = 0$.

Problem 8A.17. Suppose $T \in \mathcal{L}(V)$ is nilpotent and m is a positive integer such that $T^m = 0$.
(a) Prove that $I - T$ is invertible and that $(I - T)^{-1} = I + T + \cdots + T^{m-1}$.
(b) Explain how you would guess the formula above.

Solution. (a) With $S = I + T + \cdots + T^{m-1}$, the sum telescopes and $T^m = 0$:

$$\begin{aligned}(I - T)S &= (I + T + \cdots + T^{m-1}) - (T + T^2 + \cdots + T^m) \\ &= I - T^m = I.\end{aligned}$$

Since S is a polynomial in T, it commutes with $I - T$, so $S(I - T) = I$ too. Thus $I - T$ is invertible with $(I - T)^{-1} = I + T + \cdots + T^{m-1}$.
(b) It is the geometric series $1/(1 - x) = 1 + x + x^2 + \cdots$ with x replaced by T; nilpotence kills every term T^k with $k \geq m$, collapsing the series to a finite sum and removing any convergence question.

Problem 8A.18. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $T^{1+\dim \operatorname{range} T} = 0$.
If $\dim \operatorname{range} T < \dim V - 1$*, then this exercise improves 8.16.*

Solution. Apply 8.16 to the restriction of T to $U = \operatorname{range} T$, of dimension r. This is legitimate: U is invariant under T (as $Tu \in \operatorname{range} T = U$), and $T|_U$ is nilpotent because $T^k = 0$ for some k. If $U \neq \{0\}$, then 8.16 gives $(T|_U)^r = 0$, that is, $T^r u = 0$ for all $u \in \operatorname{range} T$, and hence for every $v \in V$,

$$T^{1+r}v = T^r(Tv) = 0.$$

If instead $U = \{0\}$, then $T = 0$ and $T^{1+r} = T = 0$.

Problem 8A.19. Suppose $T \in \mathcal{L}(V)$ is not nilpotent. Show that

$$V = \operatorname{null} T^{\dim V - 1} \oplus \operatorname{range} T^{\dim V - 1}.$$

For operators that are not nilpotent, this exercise improves 8.4.

Solution. With $n = \dim V$, the claim is 8.4, $V = \operatorname{null} T^n \oplus \operatorname{range} T^n$, once we know the $(n-1)$st and nth null spaces and ranges coincide.
For the null spaces: in the chain $\{0\} = \operatorname{null} T^0 \subseteq \cdots \subseteq \operatorname{null} T^n$ of 8.1, not all n inclusions can be strict, since strictness would give $\dim \operatorname{null} T^n \geq n$, hence $T^n = 0$ by 2.39 and T nilpotent. So $\operatorname{null} T^k = \operatorname{null} T^{k+1}$ for some $k \leq n-1$, and 8.2 propagates this:

$$\operatorname{null} T^{n-1} = \operatorname{null} T^k = \operatorname{null} T^n.$$

For the ranges: $\operatorname{range} T^n \subseteq \operatorname{range} T^{n-1}$ since $T^n v = T^{n-1}(Tv)$, and by the fundamental theorem of linear maps 3.21 with the previous display,

$$\dim \operatorname{range} T^{n-1} = n - \dim \operatorname{null} T^n = \dim \operatorname{range} T^n,$$

so the two are equal by 2.39.

Problem 8A.20. Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is normal and nilpotent. Prove that $T = 0$.

Solution. Everything follows from the claim: $T^2 v = 0 \implies Tv = 0$. Indeed, applying the normality criterion 7.20 to the vector Tv gives

$$\|T^*(Tv)\| = \|T(Tv)\| = \|T^2 v\| = 0,$$

so $T^*Tv = 0$, whence by the definition of the adjoint

$$\|Tv\|^2 = \langle Tv, Tv \rangle = \langle v, T^*Tv \rangle = 0.$$

Now fix a positive integer k and suppose $T^{k+1}v = 0$. Setting $w = T^{k-1}v$ gives $T^2 w = 0$, so $Tw = T^k v = 0$ by the claim. With the reverse inclusion from 8.1 this yields

$$\operatorname{null} T^{k+1} = \operatorname{null} T^k \quad \text{for every } k \geq 1,$$

and hence $\operatorname{null} T^m = \operatorname{null} T$ for every positive integer m. Choosing m with $T^m = 0$ (possible since T is nilpotent) gives $\operatorname{null} T = \operatorname{null} T^m = V$, that is, $T = 0$.
Method (2) (for $\mathbf{F} = \mathbf{C}$): the complex spectral theorem 7.31 diagonalizes the normal operator T in an orthonormal basis, and by 8.17(a) the only eigenvalue of a nilpotent operator is 0, so that diagonal matrix is 0.

Problem 8A.21. Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{null} T^{\dim V - 1} \neq \operatorname{null} T^{\dim V}$. Prove that T is nilpotent and that $\dim \operatorname{null} T^k = k$ for every integer k with $0 \leq k \leq \dim V$.

Solution. Write $n = \dim V$ (so $n \geq 1$). Every inclusion in the chain

$$\{0\} = \operatorname{null} T^0 \subseteq \operatorname{null} T^1 \subseteq \cdots \subseteq \operatorname{null} T^n$$

supplied by 8.1 is strict: if $\operatorname{null} T^k = \operatorname{null} T^{k+1}$ for some $k \le n-1$, then 8.2 propagates that equality forward, giving $\operatorname{null} T^{n-1} = \operatorname{null} T^n$ and contradicting the hypothesis.
Each strict inclusion raises the dimension by at least 1, so an induction on k gives $\dim \operatorname{null} T^k \ge k$ for $0 \le k \le n$; in particular $\dim \operatorname{null} T^n \ge n$, and since $\operatorname{null} T^n$ is a subspace of V this forces $\operatorname{null} T^n = V$ by 2.39. Thus $T^n = 0$ and T is nilpotent.
For the reverse inequality, the $n-k$ strict steps from $\operatorname{null} T^k$ up to $\operatorname{null} T^n$ give

$$n = \dim \operatorname{null} T^n \ge \dim \operatorname{null} T^k + (n-k),$$

so $\dim \operatorname{null} T^k \le k$. Hence $\dim \operatorname{null} T^k = k$ for every k with $0 \le k \le n$.

Problem 8A.22. Suppose $T \in \mathcal{L}(\mathbf{C}^5)$ is such that $\operatorname{range} T^4 \ne \operatorname{range} T^5$. Prove that T is nilpotent.

Solution. It suffices to show $T^5 = 0$. The ranges decrease, $\operatorname{range} T^{k+1} \subseteq \operatorname{range} T^k$ (Exercise 8A.6), and once they stabilize they stabilize forever (Exercise 8A.7); so if $\operatorname{range} T^m = \operatorname{range} T^{m+1}$ for some $m \in \{0,1,2,3\}$ we would get $\operatorname{range} T^4 = \operatorname{range} T^5$, contrary to hypothesis. Hence all five inclusions in

$$\mathbf{C}^5 = \operatorname{range} T^0 \supsetneq \operatorname{range} T^1 \supsetneq \cdots \supsetneq \operatorname{range} T^5$$

are strict, and each strict inclusion of subspaces drops the dimension by at least 1 (by 2.39), so

$$\dim \operatorname{range} T^5 \le 5 - 5 = 0.$$

Thus $\operatorname{range} T^5 = \{0\}$, that is, $T^5 = 0$.
Method (2): by Exercise 8A.9 the hypothesis is equivalent to $\operatorname{null} T^4 \ne \operatorname{null} T^5$, so Exercise 8A.21 applies.

Problem 8A.23. Give an example of an operator T on a finite-dimensional real vector space such that 0 is the only eigenvalue of T but T is not nilpotent.
[This exercise shows that (b) in 8.17 does not hold without the hypothesis that $\mathbf{F} = \mathbf{C}$.]

Solution. Take $T \in \mathcal{L}(\mathbf{R}^3)$ given by $T(x,y,z) = (-y,\, x,\, 0)$, a 90° rotation of the xy-plane killing the z-axis.
Only eigenvalue 0: the equation $T(x,y,z) = \lambda(x,y,z)$ reads

$$-y = \lambda x, \qquad x = \lambda y, \qquad 0 = \lambda z,$$

and substituting the second into the first gives $(\lambda^2+1)y = 0$, so $y = 0$ and then $x = \lambda y = 0$ since λ is real. If $\lambda \ne 0$ the third equation forces $z = 0$ as well, leaving only the zero vector; and $T(0,0,1) = 0$ shows 0 is an eigenvalue.
Not nilpotent: the minimal polynomial of T is $z(z^2+1)$, which is not of the form z^m, so 8.18 rules out nilpotence. (Directly: $T^4(1,0,0) = (1,0,0) \ne 0$, so $T^{4m} \ne 0$ for every m.)

Problem 8A.24. For each item in Example 8.15, find a basis of the domain vector space such that the matrix of the nilpotent operator with respect to that basis has the upper-triangular form promised by 8.18(c).

Solution. By 5.35 the required form says exactly that $Tv_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$ for each k, with $Tv_1 = 0$.
(a) $T(z_1,z_2,z_3,z_4) = (0,0,z_1,z_2)$, so $Te_1 = e_3$, $Te_2 = e_4$, $Te_3 = Te_4 = 0$. Reorder the standard basis as

$$v_1 = e_3, \quad v_2 = e_1, \quad v_3 = e_4, \quad v_4 = e_2,$$

so that $Tv_1 = 0$, $Tv_2 = v_1$, $Tv_3 = 0$, $Tv_4 = v_3$, and the matrix is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(b) $T \in \mathcal{L}(\mathbf{F}^3)$ has standard matrix

$$A = \begin{pmatrix} -3 & 9 & 0 \\ -7 & 9 & 6 \\ 4 & 0 & -6 \end{pmatrix},$$

and

$$A^2 = \begin{pmatrix} -54 & 54 & 54 \\ -18 & 18 & 18 \\ -36 & 36 & 36 \end{pmatrix}, \qquad A^3 = 0.$$

Take $v = (1, 0, 0)$, for which $T^2 v = (-54, -18, -36) \neq 0$ (first column of A^2), and set

$$v_1 = T^2 v = (-54, -18, -36), \quad v_2 = Tv = (-3, -7, 4), \quad v_3 = v.$$

Since $T^2 v \neq 0 = T^3 v$, Exercise 8A.2 makes $v, Tv, T^2 v$ linearly independent, hence a basis by 2.38; and $Tv_1 = 0$, $Tv_2 = v_1$, $Tv_3 = v_2$ give the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

(c) $V = \mathcal{P}_m(\mathbf{R})$, $T = D$: the standard basis $v_1 = 1, v_2 = x, \dots, v_{m+1} = x^m$ already works, since $Dv_1 = 0$ and

$$Dv_{k+1} = kx^{k-1} = kv_k \in \operatorname{span}(v_1, \dots, v_k)$$

for $k = 1, \dots, m$. The matrix has k in row k, column $k+1$ and 0 elsewhere:

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ \vdots & \vdots & \vdots & & m-1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & m \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Problem 8A.25. Suppose that V is an inner product space and $T \in \mathcal{L}(V)$ is nilpotent. Show that there is an orthonormal basis of V with respect to which the matrix of T has the upper-triangular form promised by 8.18(c).

#+begin$_{\text{m3sol}}$ Apply Gram–Schmidt to the basis furnished by 8.18(c). In detail: since T is nilpotent, 8.18 gives a basis $v_1, \dots, v_n$ of V with

$$(\star) \qquad Tv_j \in \operatorname{span}(v_1, \dots, v_{j-1}) \quad \text{for each } j,$$

this being what 5.35 says about a matrix vanishing on and below the diagonal. Gram–Schmidt (6.32) turns $v_1, \dots, v_n$ into an orthonormal list $e_1, \dots, e_n$ with

$$\operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(v_1, \dots, v_k) \quad \text{for each } k,$$

and this is an orthonormal basis by 6.28, having length $n = \dim V$.

Fix k and write $e_k = a_1 v_1 + \cdots + a_k v_k$. Applying T and using $(\star)$ together with $\operatorname{span}(v_1, \ldots, v_{j-1}) \subseteq \operatorname{span}(v_1, \ldots, v_{k-1})$ for $j \leq k$ gives

$$Te_k \in \operatorname{span}(v_1, \ldots, v_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1}),$$

both spans being $\{0\}$ when $k = 1$. So the matrix of T with respect to $e_1, \ldots, e_n$ has all entries on and below the diagonal equal to 0.

Method (2): by 8.18 the minimal polynomial of T is z^m, which splits, so 6.37 gives an orthonormal basis making $\mathcal{M}(T)$ upper triangular; its diagonal entries are the eigenvalues of T by 5.41, all 0 by 8.17(a).#+end$_{\text{m3sol}}$

8.2 Exercises 8B

Problem 8B.1. Define $T \in \mathcal{L}(\mathbf{C}^2)$ by $T(w, z) = (-z, w)$. Find the generalized eigenspaces corresponding to the distinct eigenvalues of T.

Solution. The answer is

$$G(i, T) = \operatorname{span}\big((1, -i)\big), \qquad G(-i, T) = \operatorname{span}\big((1, i)\big),$$

with $\mathbf{C}^2 = G(i, T) \oplus G(-i, T)$ as promised by 8.22.
The eigenvalues: $T(w, z) = \lambda(w, z)$ reads $-z = \lambda w$, $w = \lambda z$, which forces $-z = \lambda^2 z$ with $z \neq 0$ (else $w = \lambda z = 0$ too), so $\lambda = \pm i$; and indeed $T(1, -i) = i(1, -i)$, $T(1, i) = -i(1, i)$.
Since $\dim \mathbf{C}^2 = 2$, 8.20 gives $G(\lambda, T) = \operatorname{null}(T - \lambda I)^2$. From $T^2(w, z) = (-w, -z)$, that is $T^2 = -I$, we get

$$(T \mp iI)^2 = T^2 \mp 2iT - I = -2(I \pm iT),$$

so $G(\pm i, T) = \operatorname{null}(I \pm iT)$. Now

$$(I + iT)(w, z) = (w - iz,\ z + iw),$$

which vanishes exactly when $w = iz$, and likewise $(I - iT)(w, z) = (w + iz,\ z - iw)$ vanishes exactly when $w = -iz$. Hence $G(i, T) = \operatorname{span}((i, 1))$ and $G(-i, T) = \operatorname{span}((-i, 1))$, as claimed.

Problem 8B.2. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that $G(\lambda, T) = G\big(\frac{1}{\lambda}, T^{-1}\big)$ for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

Solution. Everything rests on the identity

$$T^{-1} - \tfrac{1}{\lambda} I = -\tfrac{1}{\lambda} T^{-1}(T - \lambda I)$$

(expand the right side). Since T^{-1} commutes with $T - \lambda I$, raising to the k-th power gives

$$\Big(T^{-1} - \tfrac{1}{\lambda} I\Big)^k = \Big(-\tfrac{1}{\lambda}\Big)^k (T^{-1})^k (T - \lambda I)^k,$$

and the prefactor $\big(-\frac{1}{\lambda}\big)^k (T^{-1})^k$ is invertible, hence injective. Therefore

$$\operatorname{null}\Big(T^{-1} - \tfrac{1}{\lambda} I\Big)^k = \operatorname{null}(T - \lambda I)^k$$

for every positive integer k. Taking $k = \dim V$ and applying 8.20 to each side gives $G(\lambda, T) = G\big(\frac{1}{\lambda}, T^{-1}\big)$.

Problem 8B.3. Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible. Prove that T and $S^{-1}TS$ have the same eigenvalues with the same multiplicities.

Solution. S maps $\operatorname{null}(S^{-1}TS - \lambda I)^k$ isomorphically onto $\operatorname{null}(T - \lambda I)^k$, for every $\lambda \in \mathbf{F}$ and every positive integer k. Indeed, $S^{-1}(\lambda I)S = \lambda I$ gives $S^{-1}TS - \lambda I = S^{-1}(T - \lambda I)S$, and the inner factors cancel on taking powers:

$$(S^{-1}TS - \lambda I)^k = S^{-1}(T - \lambda I)^k S.$$

Hence $(S^{-1}TS - \lambda I)^k v = 0$ if and only if $(T - \lambda I)^k(Sv) = 0$, since S^{-1} is injective; and S is a bijection of V, so it carries the one null space onto the other. Therefore

$$\dim \operatorname{null}(S^{-1}TS - \lambda I)^k = \dim \operatorname{null}(T - \lambda I)^k.$$

Taking $k = 1$ shows λ is an eigenvalue of $S^{-1}TS$ exactly when it is one of T; taking $k = \dim V$ and applying 8.20 to both sides gives $\dim G(\lambda, S^{-1}TS) = \dim G(\lambda, T)$, which by 8.23 is equality of multiplicities.

Problem 8B.4. Suppose $\dim V \geq 2$ and $T \in \mathcal{L}(V)$ is such that $\operatorname{null} T^{\dim V - 2} \neq \operatorname{null} T^{\dim V - 1}$. Prove that T has at most two distinct eigenvalues.

Solution. The eigenvalue 0 has multiplicity at least $n - 1$, where $n = \dim V$, which leaves room for at most one other eigenvalue.

Every inclusion in the chain $\{0\} = \operatorname{null} T^0 \subseteq \cdots \subseteq \operatorname{null} T^{n-1}$ from 8.1 is strict: the last one is the hypothesis, and an equality $\operatorname{null} T^j = \operatorname{null} T^{j+1}$ with $j \leq n-3$ would by 8.2 force $\operatorname{null} T^{n-2} = \operatorname{null} T^{n-1}$. Each of the $n - 1$ strict inclusions raises the dimension by at least 1, so using 8.1 once more and then 8.20 with $\lambda = 0$,

$$\dim G(0, T) = \dim \operatorname{null} T^n \geq \dim \operatorname{null} T^{n-1} \geq n - 1.$$

In particular $\operatorname{null} T \neq \{0\}$ (as $n \geq 2$), so 0 is an eigenvalue.

Let $\lambda_1 = 0, \lambda_2, \ldots, \lambda_m$ be the distinct eigenvalues of T. Their generalized eigenspaces sum directly: if $v_1 + \cdots + v_m = 0$ with $v_k \in G(\lambda_k, T)$, the nonzero v_k would form a linearly independent list (by 8.12) summing to 0 with unit coefficients, so all $v_k = 0$ and 1.45 applies. Hence by 3.94, and since $\dim G(\lambda_k, T) \geq \dim E(\lambda_k, T) \geq 1$ for $k \geq 2$,

$$n \geq \sum_{k=1}^{m} \dim G(\lambda_k, T) \geq (n - 1) + (m - 1),$$

so $m \leq 2$.

Problem 8B.5. Suppose $T \in \mathcal{L}(V)$ and 3 and 8 are eigenvalues of T. Let $n = \dim V$. Prove that $V = (\operatorname{null} T^{n-2}) \oplus (\operatorname{range} T^{n-2})$.

Solution. It suffices to show $\operatorname{null} T^{n-2} = \operatorname{null} T^n$ and $\operatorname{range} T^{n-2} = \operatorname{range} T^n$, since 8.4 gives $V = (\operatorname{null} T^n) \oplus (\operatorname{range} T^n)$. (Note $n \geq 2$ by 5.11, so $n - 2 \geq 0$.)

First, $\dim \operatorname{null} T^n \leq n - 2$. By 8.20, $\operatorname{null} T^n = G(0, T)$. If 0 is an eigenvalue, then $0, 3, 8$ are distinct eigenvalues, whose generalized eigenspaces sum directly (as in Exercise 8B.4, via 8.12 and 1.45), so by 3.94

$$\dim G(0, T) + \dim G(3, T) + \dim G(8, T) \leq n,$$

and the last two dimensions are each at least 1. If 0 is not an eigenvalue, then T^n is injective and $\dim \operatorname{null} T^n = 0 \leq n - 2$.

Next, let m be the least nonnegative integer with $\operatorname{null} T^m = \operatorname{null} T^{m+1}$, which exists with $m \leq n$ by 8.3. Minimality makes all m inclusions in $\{0\} \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^m$ strict (an earlier equality would propagate by 8.2), so $m \leq \dim \operatorname{null} T^m$. Since 8.2 gives $\operatorname{null} T^m = \operatorname{null} T^n$,

$$m \leq \dim \operatorname{null} T^m = \dim \operatorname{null} T^n \leq n-2,$$

so $n-2 \geq m$ and 8.2 yields $\operatorname{null} T^{n-2} = \operatorname{null} T^m = \operatorname{null} T^n$.
Finally $\operatorname{range} T^n \subseteq \operatorname{range} T^{n-2}$, and 3.21 applied to T^{n-2} and T^n gives

$$\dim \operatorname{range} T^{n-2} = n - \dim \operatorname{null} T^{n-2} = \dim \operatorname{range} T^n,$$

so the two ranges coincide by 2.39. Substituting both equalities into 8.4 finishes the proof.

Problem 8B.6. Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Explain why the exponent of $z-\lambda$ in the factorization of the minimal polynomial of T is the smallest positive integer m such that $(T-\lambda I)^m|_{G(\lambda,T)} = 0$.

Solution. Write $n = \dim V$, $G = G(\lambda, T)$, and factor the minimal polynomial as $p(z) = (z-\lambda)^k q(z)$ with $q(\lambda) \neq 0$; here $k \geq 1$ because $p(\lambda) = 0$ by 5.27(a). We show $k = m$.
The integer m exists: $G = \operatorname{null}(T-\lambda I)^n$ by 8.20, and G is invariant under T by 5.18, hence under $N := (T-\lambda I)|_G$, which satisfies $N^n = 0$. So $m \leq n$.
(i) $m \leq k$. Expanding $q(z) = c_0 + c_1(z-\lambda) + \cdots + c_d(z-\lambda)^d$ with $c_0 = q(\lambda) \neq 0$ gives

$$q(T)|_G = c_0 I|_G + A, \qquad A = c_1 N + \cdots + c_d N^d,$$

and $A = NB$ with B a polynomial in N, so $A^m = N^m B^m = 0$. An operator $c_0 I + A$ with $c_0 \neq 0$ and A nilpotent is invertible, with inverse $c_0^{-1} \sum_{i=0}^{m-1} (-c_0^{-1} A)^i$ (telescoping). Restricting $p(T) = 0$ to G thus gives

$$0 = p(T)|_G = \big((T-\lambda I)^k|_G\big)\big(q(T)|_G\big),$$

and cancelling the invertible right factor leaves $(T-\lambda I)^k|_G = 0$, so $m \leq k$ by minimality.
(ii) $k \leq m$. Put $R = \operatorname{range}(T-\lambda I)^n$, so $V = G \oplus R$ by 8.4 applied to $T-\lambda I$, and R is T-invariant by 5.18. Then $(T-\lambda I)|_R$ is injective, since $u \in R$ with $(T-\lambda I)u = 0$ lies in $G \cap R = \{0\}$; hence so is $(T-\lambda I)^k|_R$. For $u \in R$ we have $q(T)u \in R$ and $(T-\lambda I)^k(q(T)u) = p(T)u = 0$, so $q(T)|_R = 0$. Therefore $r(z) := (z-\lambda)^m q(z)$ satisfies $r(T)|_G = 0$ (definition of m) and $r(T)|_R = 0$, whence $r(T) = 0$. By 5.29 the nonzero polynomial r is a multiple of p, so

$$m + \deg q = \deg r \geq \deg p = k + \deg q.$$

Problem 8B.7. Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T with multiplicity d. Prove that $G(\lambda, T) = \operatorname{null}(T-\lambda I)^d$.
[If $d < \dim V$, then this exercise improves 8.20.]

Solution. Write $G = G(\lambda, T)$, so $d = \dim G$ by 8.23. The inclusion $\operatorname{null}(T-\lambda I)^d \subseteq G$ is immediate from the definition 8.19, so only $(T-\lambda I)^d|_G = 0$ needs proof.
By 8.20, $G = \operatorname{null}(T-\lambda I)^n$ where $n = \dim V$; hence 5.18 (with the polynomial $(z-\lambda)^n$) makes G invariant under T and so under $T-\lambda I$, and

$$N := (T-\lambda I)\big|_G \in \mathcal{L}(G)$$

satisfies $N^n = 0$. So N is a nilpotent operator on a space of dimension d, and 8.16 gives $N^d = 0$, that is, $G \subseteq \operatorname{null}(T-\lambda I)^d$.

Problem 8B.8. Suppose $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T. Prove that

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

if and only if the minimal polynomial of T equals $(z-\lambda_1)^{k_1}\cdots(z-\lambda_m)^{k_m}$ for some positive integers $k_1,\dots,k_m$.
The case $\mathbf{F}=\mathbf{C}$ *follows immediately from 5.27(b) and the generalized eigenspace decomposition (8.22); thus this exercise is interesting only when* $\mathbf{F}=\mathbf{R}$.

Solution. Write $n=\dim V$ and p for the minimal polynomial of T. Two facts are used throughout. First, the sum $G(\lambda_1,T)+\cdots+G(\lambda_m,T)$ is automatically direct, since a relation $v_1+\cdots+v_m=0$ with $v_k\in G(\lambda_k,T)$ and some $v_k\neq 0$ would make the nonzero terms a linearly dependent list of generalized eigenvectors for distinct eigenvalues, contradicting 8.12; so 1.45 applies, and the displayed decomposition is equivalent to $V=\sum_k G(\lambda_k,T)$. Second, putting $d_k=\dim G(\lambda_k,T)$, the restriction $(T-\lambda_k I)|_{G(\lambda_k,T)}$ is nilpotent (its n-th power is 0 by 8.20) and hence its d_k-th power is 0 by 8.16.
(i) Forward. Suppose $V=G(\lambda_1,T)\oplus\cdots\oplus G(\lambda_m,T)$ and set

$$q(z)=(z-\lambda_1)^{d_1}\cdots(z-\lambda_m)^{d_m}.$$

The commuting factors of $q(T)$ may be reordered to put $(T-\lambda_j I)^{d_j}$ rightmost, so $q(T)$ kills each $G(\lambda_j,T)$, hence $q(T)=0$. By 5.29, p divides q, so $p(z)=(z-\lambda_1)^{k_1}\cdots(z-\lambda_m)^{k_m}$ with $0\le k_j\le d_j$; and each $k_j\ge 1$ since each λ_j is a zero of p by 5.27(a).
(ii) Reverse. It suffices to prove, by induction on the number of distinct eigenvalues, that any $S\in\mathcal{L}(W)$ with $W\neq\{0\}$ whose minimal polynomial splits into linear factors over $\mathbf{F}$ has W equal to the sum of its generalized eigenspaces; by 5.27(a) the distinct zeros of that minimal polynomial are exactly the distinct eigenvalues of S.
If S has one distinct eigenvalue, its minimal polynomial is $(z-\mu)^k$, so $W=\operatorname{null}(S-\mu I)^k=G(\mu,S)$. Otherwise let $\mu_1,\dots,\mu_m$ be the distinct eigenvalues, $m\ge 2$, and $\ell=\dim W$. By 8.4 applied to $S-\mu_m I$, together with 8.20,

$$W=G(\mu_m,S)\oplus U,\qquad U:=\operatorname{range}(S-\mu_m I)^\ell,$$

with U invariant under S by 5.18. Here $U\neq\{0\}$: otherwise an eigenvector u for some $\mu_j\neq\mu_m$ would satisfy $0=(S-\mu_m I)^\ell u=(\mu_j-\mu_m)^\ell u$, forcing $u=0$.
Let $R=S|_U$. The minimal polynomial r of S satisfies $r(R)=r(S)|_U=0$, so by 5.29 the minimal polynomial of R divides r and therefore also splits, and by 5.27(a) every eigenvalue of R is among $\mu_1,\dots,\mu_m$. But μ_m is not one: a nonzero $u\in U$ with $Su=\mu_m u$ would lie in $\operatorname{null}(S-\mu_m I)^\ell\cap U=\{0\}$. So R has fewer than m distinct eigenvalues, and by induction U is the sum of the $G(\mu,R)$, each of which is contained in $G(\mu,S)$. Hence

$$W=G(\mu_m,S)+U\subseteq G(\mu_1,S)+\cdots+G(\mu_m,S)\subseteq W.$$

Problem 8B.9. Suppose $\mathbf{F}=\mathbf{C}$ and $T\in\mathcal{L}(V)$. Prove that there exist $D,N\in\mathcal{L}(V)$ such that $T=D+N$, the operator D is diagonalizable, N is nilpotent, and $DN=ND$.

Solution. Let $\lambda_1,\dots,\lambda_m$ be the distinct eigenvalues of T, let $n=\dim V$, and define $D\in\mathcal{L}(V)$ by

$$D(v_1+\cdots+v_m)=\lambda_1 v_1+\cdots+\lambda_m v_m,\qquad v_k\in G(\lambda_k,T);$$

this is a linear operator because $V=G(\lambda_1,T)\oplus\cdots\oplus G(\lambda_m,T)$ by 8.22, which holds since $\mathbf{F}=\mathbf{C}$. Take $N=T-D$.
D is diagonalizable: concatenating bases of the $G(\lambda_k,T)$ gives a basis of V consisting of eigenvectors of D, since $Dv=\lambda_k v$ on $G(\lambda_k,T)$; apply 5.55.
N is nilpotent: each $G(\lambda_k,T)$ is T-invariant by 8.22(a), hence N-invariant, with

$$N|_{G(\lambda_k,T)}=(T-\lambda_k I)|_{G(\lambda_k,T)},$$

nilpotent by 8.22(b); so 8.16 gives $N^n v=0$ for $v\in G(\lambda_k,T)$, and summing over k gives $N^n=0$.
$DN=ND$: for $v\in G(\lambda_k,T)$ we have $Nv\in G(\lambda_k,T)$, so $DNv=\lambda_k Nv=NDv$; now sum over k.

Problem 8B.10. Suppose V is a complex inner product space, $e_1, \ldots, e_n$ is an orthonormal basis of V, and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, each included as many times as its multiplicity. Prove that

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

See the comment after Exercise 5 in Section 7A.

Solution. Evaluate the right-hand side in a Schur basis, where it visibly dominates the diagonal. (The list $\lambda_1, \ldots, \lambda_n$ does have n entries, by 8.25.)
The right-hand side is basis-independent: for any orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$, Parseval 6.30(b) and the definition 7.1 of the adjoint give

$$\sum_k \|Te_k\|^2 = \sum_k \sum_j \big|\langle e_k, T^* f_j\rangle\big|^2 = \sum_j \|T^* f_j\|^2,$$

whose value does not involve $e_1, \ldots, e_n$.
So take the orthonormal basis $u_1, \ldots, u_n$ supplied by Schur's theorem 6.38 (available since V is a finite-dimensional complex inner product space), with respect to which $\mathcal{M}(T) = (a_{j,k})$ is upper triangular. Its diagonal lists the eigenvalues of T by 5.41, each with its multiplicity by 8.31, so $a_{1,1}, \ldots, a_{n,n}$ is a rearrangement of $\lambda_1, \ldots, \lambda_n$. Since $Tu_k = \sum_j a_{j,k} u_j$, formula 6.24 gives $\|Tu_k\|^2 = \sum_j |a_{j,k}|^2$, and discarding the off-diagonal terms yields

$$\sum_k \|Te_k\|^2 = \sum_k \sum_j |a_{j,k}|^2 \geq \sum_k |a_{k,k}|^2 = \sum_k |\lambda_k|^2.$$

Problem 8B.11. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z-7)^2(z-8)^2$.

Solution. Take

$$T(z_1, z_2, z_3, z_4) = (7z_1,\ 7z_2,\ 8z_3,\ 8z_4).$$

Its standard matrix is diagonal with entries $7, 7, 8, 8$, so by 8.31 the characteristic polynomial is $(z-7)^2(z-8)^2$.
Directly: $(T-7I)^4$ is $(z_1, z_2, z_3, z_4) \mapsto (0, 0, z_3, z_4)$ and $(T-8I)^4$ is $(z_1, z_2, z_3, z_4) \mapsto (z_1, z_2, 0, 0)$, so 8.20 gives

$$G(7, T) = \operatorname{span}(e_1, e_2), \qquad G(8, T) = \operatorname{span}(e_3, e_4),$$

each eigenvalue therefore having multiplicity 2 by 8.23.

Problem 8B.12. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z-1)(z-5)^3$ and whose minimal polynomial equals $(z-1)(z-5)^2$.

Solution. Define $T \in \mathcal{L}(\mathbf{C}^4)$ by

$$T(z_1, z_2, z_3, z_4) = (z_1,\ 5z_2 + z_3,\ 5z_3,\ 5z_4),$$

so that with respect to the standard basis e_1, e_2, e_3, e_4 the matrix of T is the block diagonal matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Explicitly, $Te_1 = e_1$, $Te_2 = 5e_2$, $Te_3 = e_2 + 5e_3$, $Te_4 = 5e_4$.

Characteristic polynomial: the matrix is upper triangular with diagonal $1, 5, 5, 5$, so 8.31 gives multiplicities 1 and 3 for the eigenvalues 1 and 5, whence $(z-1)(z-5)^3$ by 8.26.
Minimal polynomial: with $q(z) = (z-1)(z-5)^2$, the commuting factors of $q(T)$ kill each basis vector, since $(T-I)e_1 = 0$ while $(T-5I)^2$ annihilates e_2, e_3, e_4 (as $(T-5I)e_3 = e_2$ and $(T-5I)e_2 = (T-5I)e_4 = 0$). So $q(T) = 0$ and 5.29 makes the minimal polynomial p a divisor of q; by 5.27(a) its zeros are exactly 1 and 5, leaving $p = (z-1)(z-5)$ or $p = q$. The former fails:

$$(T-I)(T-5I)e_3 = (T-I)e_2 = 4e_2 \neq 0.$$

Problem 8B.13. Give an example of an operator on $\mathbf{C}^4$ whose characteristic and minimal polynomials both equal $z(z-1)^2(z-3)$.

Solution. Define $T \in \mathcal{L}(\mathbf{C}^4)$ by

$$T(z_1, z_2, z_3, z_4) = (0,\ z_2 + z_3,\ z_3,\ 3z_4).$$

On the standard basis this reads $Te_1 = 0$, $Te_2 = e_2$, $Te_3 = e_2 + e_3$, $Te_4 = 3e_4$, so the matrix of T with respect to e_1, e_2, e_3, e_4 is the block diagonal matrix

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Characteristic polynomial: the matrix is upper triangular with diagonal $0, 1, 1, 3$, so 8.31 gives multiplicities $1, 2, 1$ for the eigenvalues $0, 1, 3$, whence $z(z-1)^2(z-3)$ by 8.26.
Minimal polynomial: with $q(z) = z(z-1)^2(z-3)$, the commuting factors of $q(T)$ kill each basis vector, since $Te_1 = 0$, $(T-3I)e_4 = 0$, and $(T-I)^2$ annihilates e_2 and e_3 (as $(T-I)e_3 = e_2$ and $(T-I)e_2 = 0$). So $q(T) = 0$, and 5.29 with 5.27(a) leaves $p = z(z-1)(z-3)$ or $p = q$. The former fails:

$$T(T-I)(T-3I)e_3 = T(T-I)(e_2 - 2e_3) = T(-2e_2) = -2e_2 \neq 0.$$

Problem 8B.14. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $z(z-1)^2(z-3)$ and whose minimal polynomial equals $z(z-1)(z-3)$.

Solution. Define $T \in \mathcal{L}(\mathbf{C}^4)$ by

$$T(z_1, z_2, z_3, z_4) = (0,\ z_2,\ z_3,\ 3z_4),$$

so that the matrix of T with respect to the standard basis e_1, e_2, e_3, e_4 is the diagonal matrix with diagonal entries $0, 1, 1, 3$:

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Characteristic polynomial: the diagonal is $0, 1, 1, 3$, so 8.31 gives multiplicities $1, 2, 1$ for the eigenvalues $0, 1, 3$, whence $z(z-1)^2(z-3)$ by 8.26.
Minimal polynomial: the standard basis consists of eigenvectors, so T is diagonalizable by 5.55, and 5.62 makes the minimal polynomial the product over the distinct eigenvalues,

$$z(z-1)(z-3).$$

Problem 8B.15. Let T be the operator on $\mathbf{C}^4$ defined by $T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3)$. Find the characteristic polynomial and the minimal polynomial of T.

Solution. Both equal z^4.
T shifts the standard basis, $e_1 \mapsto e_2 \mapsto e_3 \mapsto e_4 \mapsto 0$, so $T^4 = 0$ while $T^3 e_1 = e_4 \neq 0$. Hence the minimal polynomial divides z^4 by 5.29 but not z^3, so it is z^4. Being nilpotent, T has 0 as its only eigenvalue by 8.17, and $G(0, T) = \operatorname{null} T^4 = \mathbf{C}^4$ by 8.20, so 0 has multiplicity 4 and the characteristic polynomial is z^4 by 8.26.

Problem 8B.16. Let T be the operator on $\mathbf{C}^6$ defined by

$$T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0).$$

Find the characteristic polynomial and the minimal polynomial of T.

Solution. The minimal polynomial is z^3 and the characteristic polynomial is z^6.
T acts on the standard basis by three chains, $e_1 \mapsto e_2 \mapsto e_3 \mapsto 0$, $e_4 \mapsto e_5 \mapsto 0$, $e_6 \mapsto 0$, so $T^3 = 0$ while $T^2 e_1 = e_3 \neq 0$. Hence 5.29 makes the minimal polynomial a divisor of z^3 but not of z^2, so it equals z^3. Being nilpotent, T has 0 as its only eigenvalue by 8.17, and $G(0, T) = \operatorname{null} T^6 = \mathbf{C}^6$ by 8.20, so 0 has multiplicity 6 and 8.26 gives the characteristic polynomial z^6.

Problem 8B.17. Suppose $\mathbf{F} = \mathbf{C}$ and $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the characteristic polynomial of P is $z^m(z-1)^n$, where $m = \dim \operatorname{null} P$ and $n = \dim \operatorname{range} P$.

Solution. By 8.26 it suffices to show that the eigenvalues of P lie in $\{0, 1\}$ with $\dim G(0, P) = m$ and $\dim G(1, P) = n$; the factor with exponent 0 then simply contributes nothing.
(i) Eigenvalues. If $Pv = \lambda v$ with $v \neq 0$, then $\lambda v = Pv = P^2 v = \lambda^2 v$, so $\lambda^2 = \lambda$ and $\lambda \in \{0, 1\}$.
(ii) $G(0, P) = \operatorname{null} P$. Induction gives $P^k = P$ for every $k \geq 1$, so 8.20 with $k = \dim V$ yields $G(0, P) = \operatorname{null} P^{\dim V} = \operatorname{null} P$, of dimension m.
(iii) $G(1, P) = \operatorname{range} P$. If $v = Pu$ then $Pv = P^2 u = v$, and conversely $Pv = v$ puts v in range P; hence $\operatorname{null}(P - I) = \operatorname{range} P$. Moreover $(P - I)^2 = P - 2P + I = I - P$, and since $I - P$ is also idempotent, induction gives $(P - I)^k = (-1)^k (I - P)$ for every $k \geq 1$. So by 8.20,

$$G(1, P) = \operatorname{null}(P - I)^{\dim V} = \operatorname{null}(P - I) = \operatorname{range} P,$$

of dimension n. (If $\dim V = 0$ both sides of the exercise read 1.)

Problem 8B.18. Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Explain why the following four numbers equal each other.

- (a) The exponent of $z - \lambda$ in the factorization of the minimal polynomial of T.
- (b) The smallest positive integer m such that $(T - \lambda I)^m|_{G(\lambda, T)} = 0$.
- (c) The smallest positive integer m such that $\operatorname{null}(T - \lambda I)^m = \operatorname{null}(T - \lambda I)^{m+1}$.
- (d) The smallest positive integer m such that $\operatorname{range}(T - \lambda I)^m = \operatorname{range}(T - \lambda I)^{m+1}$.

Solution. Write $N = T - \lambda I$, $n = \dim V$, p for the minimal polynomial of T, and call the four numbers a, b, c, d. We prove $d = c$, $b = c$, $a = c$.
The number c exists with $c \leq n$, since $\operatorname{null} N^n = \operatorname{null} N^{n+1}$ by 8.3; and by 8.2 the null spaces stabilize from c onward, so

$$(1) \qquad \operatorname{null} N^c = \operatorname{null} N^n = G(\lambda, T)$$

by 8.20, while minimality of c with 8.2 makes every inclusion in $\{0\} \subsetneq \operatorname{null} N^1 \subsetneq \cdots \subsetneq \operatorname{null} N^c$ strict.
$d = c$. The ranges decrease, so $\operatorname{range} N^m = \operatorname{range} N^{m+1}$ iff their dimensions agree (2.39), and 3.21 gives $\dim \operatorname{range} N^k = n - \dim \operatorname{null} N^k$; since also $\operatorname{null} N^m \subseteq \operatorname{null} N^{m+1}$, this happens iff $\operatorname{null} N^m = \operatorname{null} N^{m+1}$.

The two conditions hold for the same m, so the least such m agree.
$b = c$. Since $N^m|_{G(\lambda,T)} = 0$ means $G(\lambda, T) \subseteq \operatorname{null} N^m$, and $G(\lambda, T) = \operatorname{null} N^c$ by (1), this holds for $m \geq c$ and fails for $m < c$ by the strictness above.
$a = c$. Write $p(z) = (z - \lambda)^a q(z)$ with $q(\lambda) \neq 0$; here $a \geq 1$ by 5.27(a).
For $c \leq a$, it is enough that $\operatorname{null} N^a = \operatorname{null} N^{a+1}$. Given $v \in \operatorname{null} N^{a+1}$, the vector $w = N^a v$ satisfies $Nw = 0$, so $Tw = \lambda w$ and hence $q(T)w = q(\lambda)w$; but N^a and $q(T)$ commute, so

$$q(\lambda)w = q(T)N^a v = p(T)v = 0,$$

forcing $w = N^a v = 0$.
For $a \leq c$, suppose $a > c$ and set $r(z) = (z - \lambda)^{a-1} q(z)$, monic of degree $\deg p - 1$. For any v, $N^a(q(T)v) = p(T)v = 0$, and $\operatorname{null} N^a = \operatorname{null} N^{a-1}$ since $a - 1 \geq c$, so $r(T)v = N^{a-1}q(T)v = 0$. Thus $r(T) = 0$ with $\deg r < \deg p$, contradicting 5.22.

Problem 8B.19. Suppose $\mathbf{F} = \mathbf{C}$ and $S \in \mathcal{L}(V)$ is a unitary operator. Prove that the constant term in the characteristic polynomial of S has absolute value 1.

Solution. The constant term is $q(0)$, where by 8.26 the characteristic polynomial factors over the distinct eigenvalues $\lambda_1, \ldots, \lambda_M$ with multiplicities $d_1, \ldots, d_M$ as $q(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_M)^{d_M}$. Hence

$$|q(0)| = |-\lambda_1|^{d_1} \cdots |-\lambda_M|^{d_M} = |\lambda_1|^{d_1} \cdots |\lambda_M|^{d_M} = 1,$$

since every eigenvalue of a unitary operator has absolute value 1 by 7.54.

Problem 8B.20. Suppose that $\mathbf{F} = \mathbf{C}$ and $V_1, \ldots, V_m$ are nonzero subspaces of V such that

$$V = V_1 \oplus \cdots \oplus V_m.$$

Suppose $T \in \mathcal{L}(V)$ and each V_k is invariant under T. For each k, let p_k denote the characteristic polynomial of $T|_{V_k}$. Prove that the characteristic polynomial of T equals $p_1 \cdots p_m$.

Solution. The characteristic polynomial factors because the generalized eigenspaces do: for every $\lambda \in \mathbf{C}$,

$$G(\lambda, T) = G(\lambda, T|_{V_1}) \oplus \cdots \oplus G(\lambda, T|_{V_m}).$$

Here we set $\dim G(\lambda, R) = 0$ when λ is not an eigenvalue of R (then $R - \lambda I$ is injective, hence so is every power of it), so that 8.26 reads

$$\text{char poly of } R = \prod_{\lambda \in \mathbf{C}} (z - \lambda)^{\dim G(\lambda, R)},$$

with all but finitely many factors equal to 1.
For the claim, put $n = \dim V$ and $N = T - \lambda I$; each V_k, being invariant under T, is invariant under every power of N. If $v_k \in G(\lambda, T|_{V_k})$, then $N^{\dim V_k} v_k = 0$ by 8.20 on V_k, so $N^n v_k = 0$ as $\dim V_k \leq n$, giving $v_k \in \operatorname{null} N^n = G(\lambda, T)$ by 8.20. Conversely, if $v \in \operatorname{null} N^n$, write $v = v_1 + \cdots + v_m$ with $v_k \in V_k$; then $0 = N^n v_1 + \cdots + N^n v_m$ with $N^n v_k \in V_k$, so $N^n v_k = 0$ for each k by uniqueness of the decomposition of 0 (1.45), and

$$\operatorname{null}\bigl((T|_{V_k} - \lambda I)^n\bigr) = \operatorname{null}\bigl((T|_{V_k} - \lambda I)^{\dim V_k}\bigr) = G(\lambda, T|_{V_k})$$

since the null spaces have stopped growing by exponent $\dim V_k \leq n$ (8.3 on V_k) and by 8.20. The sum is direct because $G(\lambda, T|_{V_k}) \subseteq V_k$ and $V_1 + \cdots + V_m$ is a direct sum.

Taking dimensions (3.94) gives $\dim G(\lambda, T) = \sum_k \dim G(\lambda, T|_{V_k})$ for every $\lambda \in \mathbf{C}$, whence

$$\begin{aligned} p_1 \cdots p_m &= \prod_{k=1}^{m} \prod_{\lambda \in \mathbf{C}} (z-\lambda)^{\dim G(\lambda, T|_{V_k})} \\ &= \prod_{\lambda \in \mathbf{C}} (z-\lambda)^{\sum_{k=1}^{m} \dim G(\lambda, T|_{V_k})} \\ &= \prod_{\lambda \in \mathbf{C}} (z-\lambda)^{\dim G(\lambda, T)}, \end{aligned}$$

which is the characteristic polynomial of T. (Only finitely many factors differ from 1, so the rearrangement is legitimate.)

Problem 8B.21. Suppose $p, q \in \mathcal{P}(\mathbf{C})$ are monic polynomials with the same zeros and q is a polynomial multiple of p. Prove that there exists $T \in \mathcal{L}(\mathbf{C}^{\deg q})$ such that the characteristic polynomial of T is q and the minimal polynomial of T is p.
[This exercise implies that every monic polynomial is the characteristic polynomial of some operator.]

Solution. Take T to be block diagonal with one block $T_k = \lambda_k I + N_k$ for each distinct zero λ_k of p, as follows. Writing $\lambda_1, \ldots, \lambda_M$ for the distinct zeros of p, which are also those of q, the monic factorizations given by 4.13 are

$$p(z) = \prod_{k=1}^{M} (z-\lambda_k)^{a_k}, \qquad q(z) = \prod_{k=1}^{M} (z-\lambda_k)^{b_k},$$

with $a_k, b_k \geq 1$, and $a_k \leq b_k$ by uniqueness of the factorization in 4.13, since q is a multiple of p; put $b = b_1 + \cdots + b_M = \deg q$. (If p, q have no zeros then both equal 1 by 4.12, $\deg q = 0$, and $T = 0$ on $\mathbf{C}^0 = \{0\}$ has characteristic polynomial the empty product $1 = q$ and minimal polynomial $1 = p$.)
Define $N_k \in \mathcal{L}(\mathbf{C}^{b_k})$ on the standard basis $f_1, \ldots, f_{b_k}$ by

$$N_k f_j = f_{j+1} \text{ for } j < a_k, \qquad N_k f_j = 0 \text{ for } j \geq a_k,$$

which makes sense because $a_k \leq b_k$. Thus N_k runs the chain $f_1 \mapsto \cdots \mapsto f_{a_k} \mapsto 0$ and kills f_j for $j \geq a_k$, so $N_k^{a_k} = 0$ while $N_k^{a_k-1} f_1 = f_{a_k} \neq 0$. (Check!)
Since N_k is nilpotent, λ_k is the only eigenvalue of T_k (by 8.17(a), as μ is an eigenvalue of T_k exactly when $\mu - \lambda_k$ is one of N_k), and $G(\lambda_k, T_k) = \operatorname{null} N_k^{b_k} = \mathbf{C}^{b_k}$ by 8.20 because $b_k \geq a_k$; so λ_k has multiplicity b_k and by 8.26 the characteristic polynomial of T_k is $(z-\lambda_k)^{b_k}$. Its minimal polynomial divides $(z-\lambda_k)^{a_k}$ by 5.29, hence equals $(z-\lambda_k)^c$ with $1 \leq c \leq a_k$ (4.13), and $c < a_k$ would force $N_k^{a_k-1} = N_k^{a_k-1-c} N_k^c = 0$, false; so $c = a_k$.
Now identify $\mathbf{C}^b = V_1 \oplus \cdots \oplus V_M$, where V_k is the coordinate subspace spanned by the standard basis vectors in positions $b_1 + \cdots + b_{k-1} + 1$ through $b_1 + \cdots + b_k$, so that $V_k = \mathbf{C}^{b_k}$, and define $T \in \mathcal{L}(\mathbf{C}^b)$ by $T|_{V_k} = T_k$ (linear and well defined by uniqueness of the direct-sum decomposition). Each V_k is nonzero and invariant under T, so Exercise 20 in this section gives

$$\text{char poly of } T = (z-\lambda_1)^{b_1} \cdots (z-\lambda_M)^{b_M} = q.$$

Since T acts blockwise, $r(T)|_{V_k} = r(T_k)$ for every $r \in \mathcal{P}(\mathbf{C})$, so

$$r(T) = 0 \iff (z-\lambda_k)^{a_k} \text{ divides } r \text{ for every } k \iff p \text{ divides } r,$$

the first equivalence by 5.29 applied to each T_k and the second because the λ_k are distinct, so for $r \neq 0$ the required linear factors in the factorization 4.13 of r are disjoint. In particular $p(T) = 0$, while any nonzero polynomial annihilating T is divisible by p and hence has degree at least $\deg p$. As p is monic, p is the minimal polynomial of T (5.22).

Problem 8B.22. Suppose A and B are block diagonal matrices of the form

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}, \qquad B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix},$$

where A_k and B_k are square matrices of the same size for each $k = 1, \dots, m$. Show that AB is a block diagonal matrix of the form

$$AB = \begin{pmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{pmatrix}.$$

Solution. Index the blocks: let n_k be the common size of A_k and B_k, put $N_0 = 0$, $N_k = n_1 + \cdots + n_k$, $n = N_m$, and

$$I_k = \{N_{k-1} + 1,\ N_{k-1} + 2,\ \dots,\ N_k\},$$

so that $I_1, \dots, I_m$ partition $\{1, \dots, n\}$. Block diagonality (8.35) of A says exactly that $A_{j,l} = 0$ when $j \in I_k$ and $l \in I_{k'}$ with $k \neq k'$, while $A_{j,l} = (A_k)_{j-N_{k-1},\, l-N_{k-1}}$ when $j, l \in I_k$ (the shifted indices then lie in $\{1, \dots, n_k\}$); likewise for B.

Fix $j \in I_k$ and $l \in I_{k'}$ and expand $(AB)_{j,l} = \sum_{r=1}^{n} A_{j,r}B_{r,l}$ by 3.47.

(i) $k \neq k'$. Every summand vanishes: $A_{j,r} = 0$ for $r \notin I_k$, and for $r \in I_k$ we have $r \notin I_{k'}$, so $B_{r,l} = 0$. Hence $(AB)_{j,l} = 0$, and AB is block diagonal with block sizes $n_1, \dots, n_m$.

(ii) $k = k'$. Only the terms with $r \in I_k$ survive, so writing $r = N_{k-1} + s$,

$$\begin{aligned}(AB)_{j,l} &= \sum_{s=1}^{n_k} A_{j,\, N_{k-1}+s}\, B_{N_{k-1}+s,\, l} \\ &= \sum_{s=1}^{n_k} (A_k)_{j-N_{k-1},\, s}\, (B_k)_{s,\, l-N_{k-1}} \\ &= (A_kB_k)_{j-N_{k-1},\, l-N_{k-1}},\end{aligned}$$

the last equality being 3.47 for the n_k-by-n_k matrices A_k and B_k, whose product is defined and again n_k-by-n_k precisely because they have the same size. Thus the kth diagonal block of AB is A_kB_k, which is the asserted form.

Problem 8B.23. Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{C}$.

(a) Show that $u + iv \in G(\lambda, T_{\mathbf{C}})$ if and only if $u - iv \in G(\overline{\lambda}, T_{\mathbf{C}})$.

(b) Show that the multiplicity of λ as an eigenvalue of $T_{\mathbf{C}}$ equals the multiplicity of $\overline{\lambda}$ as an eigenvalue of $T_{\mathbf{C}}$.

(c) Use (b) and the result about the sum of the multiplicities (8.25) to show that if $\dim V$ is an odd number, then $T_{\mathbf{C}}$ has a real eigenvalue.

(d) Use (c) and the result about real eigenvalues of $T_{\mathbf{C}}$ (Exercise 17 in Section 5A) to show that if $\dim V$ is an odd number, then T has an eigenvalue (thus giving an alternative proof of 5.34).

[See Exercise 33 in Section 3B for the definition of the complexification $T_{\mathbf{C}}$.]

Solution. Everything comes from the conjugation map $\sigma : V_{\mathbf{C}} \to V_{\mathbf{C}}$ given by $\sigma(u+iv) = u-iv$, which is additive with $\sigma \circ \sigma = I$ (hence bijective, and $\sigma(0) = 0$), is conjugate-homogeneous, $\sigma(\alpha x) = \overline{\alpha}\,\sigma(x)$, and commutes with $T_{\mathbf{C}}$, since $\sigma(Tu + iTv) = Tu - iTv = T_{\mathbf{C}}\sigma(u+iv)$. (Check! For conjugate homogeneity, both $\sigma(\alpha x)$ and $\overline{\alpha}\,\sigma(x)$ equal $(au - bv) + i(-av - bu)$ when $\alpha = a + bi$ and $x = u + iv$.) Combining the three properties,

$$\begin{aligned}\sigma\big((T_{\mathbf{C}} - \lambda I)x\big) &= \sigma(T_{\mathbf{C}}x) - \sigma(\lambda x) \\ &= T_{\mathbf{C}}\sigma(x) - \overline{\lambda}\,\sigma(x) = (T_{\mathbf{C}} - \overline{\lambda} I)\sigma(x),\end{aligned}$$

using additivity of σ and $\sigma(-y) = -\sigma(y)$; induction on k then gives $\sigma\big((T_{\mathbf{C}} - \lambda I)^k x\big) = (T_{\mathbf{C}} - \overline{\lambda} I)^k \sigma(x)$ for every positive integer k.
(a) If $u + iv \in G(\lambda, T_{\mathbf{C}})$, say $(T_{\mathbf{C}} - \lambda I)^k (u + iv) = 0$ (8.19), then applying σ gives $(T_{\mathbf{C}} - \overline{\lambda} I)^k (u - iv) = \sigma(0) = 0$, so $u - iv \in G(\overline{\lambda}, T_{\mathbf{C}})$. The converse is this implication with λ replaced by $\overline{\lambda}$ and (u, v) by $(u, -v)$. Since $\sigma \circ \sigma = I$, it follows that σ restricts to a bijection of $G(\lambda, T_{\mathbf{C}})$ onto $G(\overline{\lambda}, T_{\mathbf{C}})$.
(b) Let d be the multiplicity of λ and $x_1, \ldots, x_d$ a basis of $G(\lambda, T_{\mathbf{C}})$ (8.23). Then $\sigma(x_1), \ldots, \sigma(x_d)$ is a basis of $G(\overline{\lambda}, T_{\mathbf{C}})$, so the two multiplicities agree. It lies in $G(\overline{\lambda}, T_{\mathbf{C}})$ by (a); it is independent because $\sum_j c_j \sigma(x_j) = \sigma\big(\sum_j \overline{c_j} x_j\big) = 0$ forces $\sum_j \overline{c_j} x_j = 0$ by injectivity, hence every $c_j = 0$; and it spans, because any $y \in G(\overline{\lambda}, T_{\mathbf{C}})$ has $\sigma(y) = \sum_j a_j x_j$ by (a), whence $y = \sigma(\sigma(y)) = \sum_j \overline{a_j}\, \sigma(x_j)$. (If λ is not an eigenvalue, neither is $\overline{\lambda}$: $G(\mu, T_{\mathbf{C}}) \neq \{0\}$ exactly when μ is an eigenvalue, since a composition of injective maps is injective, and σ matches the two generalized eigenspaces; both multiplicities are then 0.)
(c) Let $n = \dim V$ be odd, so $\dim_{\mathbf{C}} V_{\mathbf{C}} = n \geq 1$ and $V_{\mathbf{C}} \neq \{0\}$ (a basis of V over $\mathbf{R}$ is a basis of $V_{\mathbf{C}}$ over $\mathbf{C}$, Exercise 11 in Section 2B). By 8.25 the multiplicities of the distinct eigenvalues of $T_{\mathbf{C}}$ sum to n; write this as $S_R + S_N = n$, where S_R collects the real eigenvalues and S_N the non-real ones. By (b), conjugation permutes the eigenvalues; it fixes no non-real one and is its own inverse, so it partitions the non-real eigenvalues into pairs $\{\mu, \overline{\mu}\}$, each pair contributing $2d$ to S_N by (b). Hence S_N is even, so $S_R = n - S_N$ is odd and in particular nonzero: $T_{\mathbf{C}}$ has a real eigenvalue.
(d) By (c) some $\lambda \in \mathbf{R}$ is an eigenvalue of $T_{\mathbf{C}}$, and by Exercise 17 in Section 5A a real number is an eigenvalue of $T_{\mathbf{C}}$ if and only if it is an eigenvalue of T. Hence T has an eigenvalue.

8.3 Exercises 8C

Problem 8C.1. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator defined by $T(z_1, z_2, z_3) = (z_2, z_3, 0)$. Prove that T does not have a square root.

Solution. A square root would force $T^2 = 0$, which is false. Indeed

$$T^2(z_1, z_2, z_3) = (z_3, 0, 0), \qquad T^3(z_1, z_2, z_3) = (0, 0, 0),$$

so T is nilpotent while $T^2 \neq 0$ (because $T^2(0, 0, 1) = (1, 0, 0)$). If $R \in \mathcal{L}(\mathbf{C}^3)$ satisfied $R^2 = T$, then $R^6 = T^3 = 0$, so R is nilpotent and hence $R^3 = 0$ by 8.16 (as $\dim \mathbf{C}^3 = 3$); therefore

$$T^2 = R^4 = R^3 R = 0,$$

a contradiction.

Problem 8C.2. Define $T \in \mathcal{L}(\mathbf{F}^5)$ by $T(x_1, x_2, x_3, x_4, x_5) = (2x_2, 3x_3, -x_4, 4x_5, 0)$.
(a) Show that T is nilpotent.
(b) Find a square root of $I + T$.

Solution. (a) $T^5 = 0$. On the standard basis, $Te_1 = 0$ and $Te_2 = 2e_1$, $Te_3 = 3e_2$, $Te_4 = -e_3$, $Te_5 = 4e_4$, so T maps $\operatorname{span}(e_1, \ldots, e_k)$ into $\operatorname{span}(e_1, \ldots, e_{k-1})$; five applications send $\mathbf{F}^5$ to $\{0\}$. Explicitly,

$$\begin{aligned} T^2(x_1, \ldots, x_5) &= (6x_3, -3x_4, -4x_5, 0, 0), \\ T^3(x_1, \ldots, x_5) &= (-6x_4, -12x_5, 0, 0, 0), \\ T^4(x_1, \ldots, x_5) &= (-24x_5, 0, 0, 0, 0), \\ T^5(x_1, \ldots, x_5) &= (0, 0, 0, 0, 0). \end{aligned}$$

(b) A square root of $I + T$ is

$$R = I + \tfrac{1}{2}T - \tfrac{1}{8}T^2 + \tfrac{1}{16}T^3 - \tfrac{5}{128}T^4.$$

Following 8.39, look for $R = I + a_1T + a_2T^2 + a_3T^3 + a_4T^4$; squaring and using $T^k = 0$ for $k \geq 5$,

$$\begin{aligned} R^2 = I &+ 2a_1T + (2a_2 + a_1^2)T^2 + (2a_3 + 2a_1a_2)T^3 \\ &+ (2a_4 + 2a_1a_3 + a_2^2)T^4. \end{aligned}$$

Setting the four coefficients equal to $1, 0, 0, 0$ and solving successively gives $a_1 = \frac{1}{2}$, $a_2 = -\frac{1}{8}$, $a_3 = \frac{1}{16}$, $a_4 = -\frac{5}{128}$ (Check!), so $R^2 = I + T$; the coefficients are real, so this works over $\mathbf{R}$ as well as $\mathbf{C}$. By the powers computed in (a), the matrix of R with respect to the standard basis is

$$\begin{pmatrix} 1 & 1 & -\frac{3}{4} & -\frac{3}{8} & \frac{15}{16} \\ 0 & 1 & \frac{3}{2} & \frac{3}{8} & -\frac{3}{4} \\ 0 & 0 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

whose square is the matrix of $I + T$, with 1's on the diagonal and $2, 3, -1, 4$ directly above it.

Problem 8C.3. Suppose V is a complex vector space. Prove that every invertible operator on V has a cube root.

Solution. Imitate 8.39 and 8.41 with cubes in place of squares.
(i) If $N \in \mathcal{L}(V)$ is nilpotent, then $I + N$ has a cube root. Fix s with $N^s = 0$; if $s = 1$ take I, so assume $s \geq 2$. For $p(x) = b_0 + b_1x + \cdots + b_{s-1}x^{s-1}$ with $b_0 = 1$, the coefficient of x^k in $p(x)^3$ is

$$c_k = \sum_{\substack{i+j+\ell=k \\ i,j,\ell \geq 0}} b_ib_jb_\ell = 3b_k + q_k(b_1, \ldots, b_{k-1}) \quad (k \geq 1),$$

because the only triples with an index equal to k are $(k, 0, 0)$, $(0, k, 0)$, $(0, 0, k)$, each contributing $b_kb_0^2 = b_k$, while every other triple has all indices at most $k - 1$; here q_k is a polynomial in the earlier coefficients, $q_1 = 0$, and $c_0 = 1$. So set $b_1 = \frac{1}{3}$ and, recursively, $b_k = -\frac{1}{3}q_k(b_1, \ldots, b_{k-1})$ for $2 \leq k \leq s - 1$, making $c_1 = 1$ and $c_k = 0$ thereafter. Then $p(x)^3 = 1 + x +$ (terms of degree at least s), so $R = p(N)$ satisfies $R^3 = I + N$, since $N^k = 0$ for $k \geq s$.
(ii) Every $\lambda = r(\cos\theta + i\sin\theta) \in \mathbf{C}$ with $r \geq 0$ has the cube root $r^{1/3}\big(\cos\frac{\theta}{3} + i\sin\frac{\theta}{3}\big)$, by the addition formulas applied twice.
Now let $T \in \mathcal{L}(V)$ be invertible and $\lambda_1, \ldots, \lambda_m$ its distinct eigenvalues, of which there is at least one by 5.19 (V is nonzero, finite-dimensional, complex). By 8.22,

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T),$$

each $G(\lambda_k, T)$ is invariant under T, and $T|_{G(\lambda_k,T)} = \lambda_kI + N_k$ with N_k nilpotent. Invertibility makes 0 a non-eigenvalue, so $\lambda_k \neq 0$ and $T|_{G(\lambda_k,T)} = \lambda_k(I + N_k/\lambda_k)$ with N_k/λ_k nilpotent. Choose S_k with $S_k^3 = I + N_k/\lambda_k$ by (i) and c_k with $c_k^3 = \lambda_k$ by (ii), and set $R_k = c_kS_k$, so $R_k^3 = T|_{G(\lambda_k,T)}$. Define $R \in \mathcal{L}(V)$ by $Rv = R_1u_1 + \cdots + R_mu_m$ where $v = u_1 + \cdots + u_m$ is the (unique, hence linearly obtained) decomposition with $u_k \in G(\lambda_k, T)$. Since $R_ku_k \in G(\lambda_k, T)$, applying R three times gives

$$R^3v = R_1^3u_1 + \cdots + R_m^3u_m = Tu_1 + \cdots + Tu_m = Tv.$$

Hence $R^3 = T$.

Problem 8C.4. Suppose V is a real vector space. Prove that the operator $-I$ on V has a square root if and only if $\dim V$ is an even number.

Solution. If $R \in \mathcal{L}(V)$ satisfies $R^2 = -I$, then R has no eigenvalue: $Rv = \lambda v$ with $v \neq 0$ and $\lambda \in \mathbf{R}$ would give $\lambda^2v = R^2v = -v$, hence $\lambda^2 = -1$, impossible over $\mathbf{R}$. Were $\dim V$ odd, R would have an

eigenvalue by 5.34; hence $\dim V$ is even.
Conversely, suppose $\dim V = 2n$ with $n \geq 1$ ($V \neq \{0\}$ throughout this chapter), let $v_1, \ldots, v_{2n}$ be a basis of V, and define $R \in \mathcal{L}(V)$ on that basis by 3.4:

$$Rv_{2k-1} = v_{2k}, \qquad Rv_{2k} = -v_{2k-1} \qquad \text{for } k = 1, \ldots, n.$$

Then $R^2 v_{2k-1} = -v_{2k-1}$ and $R^2 v_{2k} = -v_{2k}$, so R^2 agrees with $-I$ on a basis of V and therefore $R^2 = -I$.

Problem 8C.5. Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is the operator defined by $T(w, z) = (-w - z,\ 9w + 5z)$. Find a Jordan basis for T.

Solution. A Jordan basis is $v_1 = (-3, 9)$, $v_2 = (1, 0)$.
The only eigenvalue of T is 2. Indeed, $T(w, z) = \lambda(w, z)$ with $(w, z) \neq (0, 0)$ reads $-w - z = \lambda w$ and $9w + 5z = \lambda z$; here $w \neq 0$ (else the first equation gives $z = 0$), so $z = -(\lambda + 1)w$, and substituting into the second and dividing by w gives

$$9 - 5(\lambda + 1) = -\lambda(\lambda + 1), \qquad \text{that is,} \qquad (\lambda - 2)^2 = 0.$$

The eigenvectors are the nonzero multiples of $(1, -3)$, so $\dim E(2, T) = 1 < 2$ and no basis of eigenvectors exists; since a Jordan matrix is upper triangular with the eigenvalues on its diagonal (5.41), a Jordan matrix for T must be a single 2-by-2 block with 2 on the diagonal.
Take $v_2 = (1, 0)$ and $v_1 = (T - 2I)v_2 = (-3, 9) \neq 0$. Then

$$Tv_1 = (-6, 18) = 2v_1, \qquad Tv_2 = (-1, 9) = v_1 + 2v_2,$$

and v_1, v_2 is linearly independent (the second coordinate of $av_1 + bv_2 = 0$ forces $a = 0$, then $b = 0$), hence a basis by 2.38. Therefore

$$\mathcal{M}\big(T, (v_1, v_2)\big) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},$$

a single Jordan block.

Problem 8C.6. Find a basis of $\mathcal{P}_4(\mathbf{R})$ that is a Jordan basis for the differentiation operator D on $\mathcal{P}_4(\mathbf{R})$ defined by $Dp = p'$.

Solution. A Jordan basis for D is

$$p_1 = 1, \quad p_2 = x, \quad p_3 = \frac{x^2}{2}, \quad p_4 = \frac{x^3}{6}, \quad p_5 = \frac{x^4}{24},$$

that is, $p_k = x^{k-1}/(k-1)!$; this is the list D^4u, D^3u, D^2u, Du, u of 8.42 with $u = x^4/24$. Its entries are nonzero scalar multiples of the standard basis $1, x, x^2, x^3, x^4$, hence a basis of $\mathcal{P}_4(\mathbf{R})$, and

$$Dp_1 = 0, \quad Dp_2 = p_1, \quad Dp_3 = p_2, \quad Dp_4 = p_3, \quad Dp_5 = p_4,$$

so the matrix of D with respect to $p_1, \ldots, p_5$ is the single Jordan block

$$\mathcal{M}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The diagonal entry 0 is forced: $D^5 = 0$ by 8.15(c), so D is nilpotent with 0 as its only eigenvalue (8.17), and a Jordan matrix is upper triangular with the eigenvalues on its diagonal (5.41).

Problem 8C.7. Suppose $T \in \mathcal{L}(V)$ is nilpotent and $v_1, \ldots, v_n$ is a Jordan basis for T. Prove that the minimal polynomial of T is z^{m+1}, where m is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the matrix of T with respect to $v_1, \ldots, v_n$.

Solution. The minimal polynomial is z^d, where $d = \max\{d_1, \ldots, d_p\}$ is the largest Jordan block size, and $d = m + 1$.

By 8.44 the matrix of T is block diagonal with blocks $A_1, \ldots, A_p$ of sizes $d_1, \ldots, d_p$ summing to n, the kth carrying a scalar λ_k on its diagonal and 1's directly above it. That matrix is upper triangular, so its diagonal entries are the eigenvalues of T (5.41), all equal to 0 since T is nilpotent (8.17). Writing $u_{k,1}, \ldots, u_{k,d_k}$ for the basis vectors of the kth block, the columns give $Tu_{k,1} = 0$ and $Tu_{k,i} = u_{k,i-1}$ for $2 \le i \le d_k$, whence by induction on j

$$T^j u_{k,i} = \begin{cases} u_{k,\,i-j} & \text{if } 0 \le j \le i-1, \\ 0 & \text{if } j \ge i. \end{cases}$$

(i) $m = d - 1$. The entry in position $(i, i+1)$ is 1 when i and $i+1$ lie in the same block and 0 when i ends one block and $i+1$ begins the next, so the line above the diagonal reads

$$\underbrace{1, \ldots, 1}_{d_1 - 1},\ 0,\ \underbrace{1, \ldots, 1}_{d_2 - 1},\ 0,\ \ldots,\ 0,\ \underbrace{1, \ldots, 1}_{d_p - 1},$$

whose consecutive strings of 1's have lengths $d_1 - 1, \ldots, d_p - 1$; hence $m = d - 1$ (also when every $d_k = 1$, so that no 1's appear and $m = 0$).

(ii) $T^d = 0$ and $T^{d-1} \neq 0$. The displayed formula gives $T^d u_{k,i} = 0$ for all k, i, since $d \ge d_k \ge i$, and T^d kills a basis. Choosing k with $d_k = d$: if $d \ge 2$ then $T^{d-1} u_{k,d} = u_{k,1} \neq 0$, and if $d = 1$ then $T^{d-1} = I \neq 0$ as $V \neq \{0\}$.

Now let $q(z) = c_0 + c_1 z + \cdots + c_{e-1} z^{e-1} + z^e$ be the minimal polynomial of T (5.24). Minimality of the degree and $T^d = 0$ give $e \le d$. If $e < d$, choose k with $d_k = d$, set $c_e = 1$, and apply $q(T)$ to $u_{k,d}$:

$$0 = q(T) u_{k,d} = \sum_{i=0}^{e} c_i T^i u_{k,d} = \sum_{i=0}^{e} c_i\, u_{k,\,d-i},$$

where $d - i \ge d - e \ge 1$ for each i, so the $u_{k,d-i}$ are distinct members of the basis and every $c_i = 0$, contradicting $c_e = 1$. Hence $e = d$, and since z^d is monic of degree d with $T^d = 0$, uniqueness in 5.24 gives $q(z) = z^d = z^{m+1}$.

Problem 8C.8. Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T^2 with respect to this basis.

Solution. Writing $\mathcal{M}(T)$ as block diagonal with blocks $A_1, \ldots, A_p$ of sizes $d_1, \ldots, d_p$, where A_k carries a scalar λ_k on the diagonal and 1 directly above it (8.44), the matrix $\mathcal{M}(T^2)$ is block diagonal with the same block sizes, its kth block being upper triangular with λ_k^2 on the diagonal, $2\lambda_k$ on the first superdiagonal, 1 on the second, and 0 elsewhere. For $d_k = 5$, say,

$$A_k^2 = \begin{pmatrix} \lambda_k^2 & 2\lambda_k & 1 & 0 & 0 \\ 0 & \lambda_k^2 & 2\lambda_k & 1 & 0 \\ 0 & 0 & \lambda_k^2 & 2\lambda_k & 1 \\ 0 & 0 & 0 & \lambda_k^2 & 2\lambda_k \\ 0 & 0 & 0 & 0 & \lambda_k^2 \end{pmatrix}.$$

Indeed $\mathcal{M}(T^2) = \mathcal{M}(T)^2$ by 3.81, and squaring a block diagonal matrix squares each block (Exercise 22 in Section 8B), so the kth block of $\mathcal{M}(T^2)$ is A_k^2. Writing $A_k = \lambda_k I + N_k$ with N_k carrying 1 on the first superdiagonal and 0 elsewhere, and using that I commutes with N_k,

$$A_k^2 = (\lambda_k I + N_k)^2 = \lambda_k^2 I + 2\lambda_k N_k + N_k^2,$$

where N_k^2 has 1 on the second superdiagonal and 0 elsewhere (and $N_k^2 = 0$ when $d_k \leq 2$). (Check!)

Problem 8C.9. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Explain why there exist $v_1, \ldots, v_n \in V$ and nonnegative integers $m_1, \ldots, m_n$ such that (a) and (b) below both hold.

- (a) $T^{m_1}v_1, \ldots, Tv_1, v_1, \ldots, T^{m_n}v_n, \ldots, Tv_n, v_n$ is a basis of V.
- (b) $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$.

Solution. Take a Jordan basis for T, which exists by 8.45, write $u_1^k, \ldots, u_{d_k}^k$ for its vectors belonging to the kth block, and set

$$n = p, \qquad v_k = u_{d_k}^k, \qquad m_k = d_k - 1 \geq 0.$$

By 8.44 the matrix of T with respect to this basis is block diagonal with blocks $A_1, \ldots, A_p$, the kth being d_k-by-d_k with a scalar λ_k on the diagonal, 1 directly above it, and 0 elsewhere. That matrix is upper triangular, so its diagonal entries are the eigenvalues of T (5.41), all equal to 0 since T is nilpotent (8.17). Reading off the jth column of A_k (which is 0 for $j = 1$, and for $j \geq 2$ has a single 1, in row $j-1$),

$$Tu_1^k = 0 \qquad \text{and} \qquad Tu_j^k = u_{j-1}^k \quad \text{for } j = 2, \ldots, d_k,$$

so induction on i gives $T^i v_k = u_{d_k - i}^k$ for $i = 0, 1, \ldots, d_k - 1$.
This yields (b), since $T^{m_k+1}v_k = T\big(T^{d_k-1}v_k\big) = Tu_1^k = 0$, and (a), since the same formula makes $T^{m_k}v_k, \ldots, Tv_k, v_k$ equal to $u_1^k, u_2^k, \ldots, u_{d_k}^k$, the kth block of basis vectors in its original order; concatenating over k recovers the Jordan basis itself.

Problem 8C.10. Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis $v_n, \ldots, v_1$ obtained by reversing the order of the v's.

Solution. Reversing the basis puts the Jordan blocks in reverse order and transposes each one:

$$\mathcal{M}\big(T, (v_n, \ldots, v_1)\big) = \begin{pmatrix} A_p^{\mathrm{t}} & & 0 \\ & \ddots & \\ 0 & & A_1^{\mathrm{t}} \end{pmatrix}, \qquad A_k^{\mathrm{t}} = \begin{pmatrix} \lambda_k & & & 0 \\ 1 & \ddots & & \\ & \ddots & \ddots & \\ 0 & & 1 & \lambda_k \end{pmatrix},$$

where $A = \mathcal{M}\big(T, (v_1, \ldots, v_n)\big)$ is block diagonal with blocks $A_1, \ldots, A_p$ of sizes $d_1, \ldots, d_p$, the kth carrying λ_k on the diagonal and 1 directly above it (8.44). So the reversed matrix is lower triangular, each block having an eigenvalue of T repeated on its diagonal and 1's directly below it.
Reversing a basis rotates its matrix through 180 degrees. Indeed, with $w_j = v_{n+1-j}$ and $B = \mathcal{M}\big(T, (w_1, \ldots, w_n)\big)$, substituting $j = n + 1 - i$ gives

$$Tw_k = \sum_{i=1}^{n} A_{i,\, n+1-k}\, w_{n+1-i} = \sum_{j=1}^{n} A_{n+1-j,\, n+1-k}\, w_j,$$

so $B_{j,k} = A_{n+1-j,\, n+1-k}$. Rotating a block diagonal matrix reverses the order of its blocks and rotates each of them; rotating A_k fixes its diagonal and carries the line directly above the diagonal to the line directly below, which is exactly A_k^{t}.

Problem 8C.11. Suppose $T \in \mathcal{L}(V)$. Explain why every vector in each Jordan basis for T is a generalized eigenvector of T.

Solution. The jth vector of a Jordan block with diagonal entry λ_k is nonzero and killed by $(T - \lambda_k I)^j$, so 8.8 makes it a generalized eigenvector corresponding to λ_k.

In detail: by 8.44 the matrix of a Jordan basis $v_1, \dots, v_n$ is block diagonal with blocks $A_1, \dots, A_p$, the kth being d_k-by-d_k with a scalar λ_k on the diagonal, 1 directly above it, and 0 elsewhere, and each v_j belongs to exactly one block. Writing $u_1, \dots, u_{d_k}$ for the vectors of the kth block, the columns of the block diagonal matrix give

$$(T - \lambda_k I)u_j = u_{j-1} \quad \text{for } j = 1, \dots, d_k, \qquad u_0 = 0.$$

Iterating, $(T - \lambda_k I)^i u_j = u_{j-i}$ for $0 \le i \le j$, so $i = j$ gives $(T - \lambda_k I)^j u_j = 0$. Here $u_j \neq 0$, being a member of a basis, and λ_k is an eigenvalue of T, since $(T - \lambda_k I)u_1 = 0$ with $u_1 \neq 0$.

Problem 8C.12. Suppose $T \in \mathcal{L}(V)$ is diagonalizable. Show that $\mathcal{M}(T)$ is a diagonal matrix with respect to every Jordan basis for T.

Solution. A Jordan block of size at least 2 would produce a generalized eigenvector that is not an eigenvector, which a diagonalizable operator does not have; so every block has size 1.
Let $v_1, \dots, v_n$ be any Jordan basis for T, with blocks $A_1, \dots, A_p$ of sizes $d_1, \dots, d_p$ summing to n, the kth carrying λ_k on the diagonal and 1 directly above it (8.44). Then $\mathcal{M}(T)$ is diagonal exactly when every $d_k = 1$, since a block of size at least 2 contributes a 1 off the diagonal.
Claim: if T is diagonalizable, $\lambda \in \mathbf{F}$, and $(T - \lambda I)^2 v = 0$, then $(T - \lambda I)v = 0$. Indeed, letting $\lambda'_1, \dots, \lambda'_m$ be the distinct eigenvalues of T, diagonalizability gives $V = E(\lambda'_1, T) \oplus \cdots \oplus E(\lambda'_m, T)$ by 5.55; write $v = x_1 + \cdots + x_m$ with $x_i \in E(\lambda'_i, T)$. Each eigenspace is invariant under $T - \lambda I$, which acts on it as multiplication by $\lambda'_i - \lambda$, so

$$0 = (T - \lambda I)^2 v = (\lambda'_1 - \lambda)^2 x_1 + \cdots + (\lambda'_m - \lambda)^2 x_m,$$

whose ith summand lies in $E(\lambda'_i, T)$; directness of the sum forces every summand to vanish (1.45), so $x_i = 0$ whenever $\lambda'_i \neq \lambda$. Hence every term of $(T - \lambda I)v = \sum_i (\lambda'_i - \lambda)x_i$ vanishes.
Now if some $d_k \geq 2$, the first two columns of A_k give $Tu_1 = \lambda_k u_1$ and $Tu_2 = \lambda_k u_2 + u_1$ for the corresponding basis vectors, that is,

$$(T - \lambda_k I)u_2 = u_1 \qquad \text{and} \qquad (T - \lambda_k I)^2 u_2 = 0,$$

so the claim (with $\lambda = \lambda_k$, $v = u_2$) yields $u_1 = (T - \lambda_k I)u_2 = 0$, impossible for a member of a basis. Hence every $d_k = 1$ and $\mathcal{M}(T)$ is diagonal, for every Jordan basis.

Problem 8C.13. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that if $v_1, \dots, v_n$ are vectors in V and $m_1, \dots, m_n$ are nonnegative integers such that

$$T^{m_1} v_1, \dots, T v_1, v_1, \dots, T^{m_n} v_n, \dots, T v_n, v_n \quad \text{is a basis of } V$$

and

$$T^{m_1 + 1} v_1 = \cdots = T^{m_n + 1} v_n = 0,$$

then $T^{m_1} v_1, \dots, T^{m_n} v_n$ is a basis of $\operatorname{null} T$.
[This exercise shows that $n = \dim \operatorname{null} T$. Thus the positive integer n that appears above depends only on T and not on the specific Jordan basis chosen for T.]

Solution. The list $T^{m_1} v_1, \dots, T^{m_n} v_n$ is a sublist of the given basis $\mathcal{B}$, lies in $\operatorname{null} T$, and spans it.
Its members are the n distinct vectors $T^i v_j$ of $\mathcal{B}$ with $i = m_j$, so the list is linearly independent, every sublist of a linearly independent list being linearly independent; and it lies in $\operatorname{null} T$ because $T\big(T^{m_j} v_j\big) = T^{m_j + 1} v_j = 0$ by hypothesis.
For spanning, take $v \in \operatorname{null} T$ and write $v = \sum_{j=1}^{n} \sum_{i=0}^{m_j} a_{j,i} T^i v_j$ in the basis $\mathcal{B}$. Applying T, the $i = m_j$

terms drop out since $T^{m_j+1}v_j = 0$, leaving

$$0 = Tv = \sum_{j=1}^{n} \sum_{i=0}^{m_j-1} a_{j,i} T^{i+1} v_j.$$

Each $T^{i+1}v_j$ occurring here has $1 \le i+1 \le m_j$, hence is a member of $\mathcal{B}$, and distinct pairs (j,i) give distinct members; linear independence of $\mathcal{B}$ therefore forces $a_{j,i} = 0$ for all $i \le m_j - 1$, so that

$$v = \sum_{j=1}^{n} a_{j,m_j} T^{m_j} v_j \in \operatorname{span}\big(T^{m_1}v_1, \dots, T^{m_n}v_n\big).$$

Problem 8C.14. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that there does not exist a direct sum decomposition of V into two nonzero subspaces invariant under T if and only if the minimal polynomial of T is of the form $(z-\lambda)^{\dim V}$ for some $\lambda \in \mathbf{C}$.

Solution. Both directions turn on the number of Jordan blocks: no such decomposition exists exactly when T has a single Jordan block, of size $n = \dim V \ge 1$. Both use the fact that a monic $q \in \mathcal{P}(\mathbf{C})$ dividing $(z-\lambda)^N$ equals $(z-\lambda)^r$ with $r = \deg q \le N$: a nonconstant such q factors as $(z-\mu_1)\cdots(z-\mu_r)$ by 4.13, and each μ_i, being a zero of q, is a zero of $(z-\lambda)^N$, so $\mu_i = \lambda$.
(i) Suppose no such decomposition exists. Since $\mathbf{F} = \mathbf{C}$, 8.46 supplies a Jordan basis $v_1, \dots, v_n$, with blocks $A_1, \dots, A_p$ of sizes $d_1, \dots, d_p$ as in 8.44; let U_k be the span of the basis vectors of the kth block. Reading off the columns, $Tu_j = \lambda_k u_j + u_{j-1} \in U_k$ (with $u_0 = 0$), so each U_k is nonzero and invariant under T, and $V = U_1 \oplus \cdots \oplus U_p$ since concatenating their bases returns $v_1, \dots, v_n$. If $p \ge 2$, then $U = U_1$ and $W = U_2 + \cdots + U_p$ are nonzero, invariant under T (a sum of invariant subspaces is invariant), and satisfy $U + W = V$ with $\dim U + \dim W = n$, so $V = U \oplus W$ (2.39 with 3.94), contrary to hypothesis. Hence $p = 1$: one block of size n, with some $\lambda \in \mathbf{C}$ on its diagonal, so $(T - \lambda I)v_j = v_{j-1}$ for $j = 1, \dots, n$ with $v_0 = 0$. Iterating gives $(T-\lambda I)^i v_j = v_{j-i}$ for $0 \le i \le j$, so $(T-\lambda I)^n = 0$ while $(T-\lambda I)^{n-1}v_n = v_1 \ne 0$. The minimal polynomial q of T therefore divides $(z-\lambda)^n$ by 5.29, so $q = (z-\lambda)^r$ with $r \le n$, and $r \le n-1$ is ruled out by $(T-\lambda I)^{n-1} \ne 0$. Thus $q = (z-\lambda)^{\dim V}$.
(ii) Conversely, let the minimal polynomial of T be $(z-\lambda)^n$ and suppose $V = U \oplus W$ with U, W nonzero and invariant under T. The minimal polynomial p_U of $T|_U$ divides $(z-\lambda)^n$ by 5.31, hence $p_U = (z-\lambda)^a$ with $a = \deg p_U \le \dim U$ by 5.22; likewise $T|_W$ has minimal polynomial $(z-\lambda)^b$ with $b \le \dim W$. Put $c = \max\{a,b\}$. Then $(z-\lambda)^c$ is a multiple of both, so 5.29 applied to $T|_U$ and to $T|_W$ gives $(T-\lambda I)^c = 0$ on U and on W, hence on $V = U + W$; by 5.29 again $(z-\lambda)^n$ divides $(z-\lambda)^c$, so $n \le c$. But U and W are nonzero with $\dim U + \dim W = n$ (3.94), so

$$c = \max\{a,b\} \le \max\{\dim U, \dim W\} \le n-1,$$

a contradiction.

8.4 Exercises 8D

Problem 8D.1. Suppose V is an inner product space and $v, w \in V$. Define an operator $T \in \mathcal{L}(V)$ by $Tu = \langle u, v\rangle w$. Find a formula for $\operatorname{tr} T$.

Solution.

$$\operatorname{tr} T = \langle w, v\rangle.$$

Take an orthonormal basis $e_1, \dots, e_n$ of V, which exists by 6.35 since V is finite-dimensional. Computing the trace from it (8.55) and pulling scalars out of the first slot,

$$\operatorname{tr} T = \sum_{k=1}^{n} \langle Te_k, e_k\rangle = \sum_{k=1}^{n} \langle e_k, v\rangle \langle w, e_k\rangle,$$

which is exactly $\langle w, v\rangle$, obtained by taking the inner product with v of the expansion $w = \langle w, e_1\rangle e_1 + \cdots + \langle w, e_n\rangle e_n$ (6.30(a)).

Problem 8D.2. Suppose $P \in \mathcal{L}(V)$ satisfies $P^2 = P$. Prove that

$$\operatorname{tr} P = \dim \operatorname{range} P.$$

Solution. With respect to a basis adapted to $V = \operatorname{range} P \oplus \operatorname{null} P$, the matrix of P is diagonal with $\dim \operatorname{range} P$ entries 1 and the rest 0, so its trace is $\dim \operatorname{range} P$.
First, $\operatorname{range} P = \{v \in V : Pv = v\}$: if $v = Pu$, then $Pv = P^2u = Pu = v$, and the reverse inclusion is immediate. Hence $\operatorname{null} P \cap \operatorname{range} P = \{0\}$, since such a v satisfies $v = Pv = 0$; and every $v \in V$ splits as $v = Pv + (v - Pv)$, where $Pv \in \operatorname{range} P$ and $P(v - Pv) = Pv - P^2v = 0$.
Now let $u_1, \ldots, u_m$ be a basis of $\operatorname{range} P$ and $w_1, \ldots, w_k$ a basis of $\operatorname{null} P$ (either list may be empty). The concatenated list spans V by the previous paragraph, and its length $m + k$ equals $\dim V$ by 3.21, so it is a basis of V (2.42). On it,

$$Pu_j = u_j \quad \text{for } j = 1, \ldots, m, \qquad Pw_j = 0 \quad \text{for } j = 1, \ldots, k,$$

so $\mathcal{M}(P)$ is diagonal with m entries 1 then k entries 0. Since the trace may be computed from the matrix with respect to any basis (8.50, 8.51), $\operatorname{tr} P = m = \dim \operatorname{range} P$.

Problem 8D.3. Suppose $T \in \mathcal{L}(V)$ and $T^5 = T$. Prove that the real and imaginary parts of $\operatorname{tr} T$ are both integers.

Solution. Every eigenvalue λ of T satisfies $\lambda^5 = \lambda$, so over $\mathbf{C}$ the trace is an integer combination of $0, 1, -1, i, -i$.
(i) $\mathbf{F} = \mathbf{C}$. If $Tv = \lambda v$ with $v \neq 0$, then $\lambda^5 v = T^5 v = Tv = \lambda v$, so $(\lambda^5 - \lambda)v = 0$ and hence

$$\lambda(\lambda - 1)(\lambda + 1)(\lambda^2 + 1) = 0, \qquad \lambda \in \{0,\ 1,\ -1,\ i,\ -i\}.$$

By 8.52, $\operatorname{tr} T = d_1\lambda_1 + \cdots + d_m\lambda_m$, where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T and the multiplicities d_j are positive integers (8.23). Letting a be the sum of the d_j with $\lambda_j = 1$ minus the sum of those with $\lambda_j = -1$, and b the corresponding difference for i and $-i$ (the terms with $\lambda_j = 0$ contributing nothing), we get $\operatorname{tr} T = a + bi$ with a, b integers.
(ii) $\mathbf{F} = \mathbf{R}$. Let $A = \mathcal{M}(T)$ with respect to any basis of V, an n-by-n real matrix; then $A^5 = \mathcal{M}(T^5) = A$ by 3.43. Let $S \in \mathcal{L}(\mathbf{C}^n)$ be the operator whose matrix with respect to the standard basis is A; then $\mathcal{M}(S^5) = A^5 = \mathcal{M}(S)$ by 3.43 again, so $S^5 = S$, an operator being determined by its matrix with respect to a fixed basis. By (i) the real and imaginary parts of $\operatorname{tr} S$ are integers, and $\operatorname{tr} S = \operatorname{tr} T$, both being the sum of the diagonal entries of A. Since $\operatorname{tr} T$ is real, its imaginary part is 0 and its real part is the integer $\operatorname{tr} S$.

Problem 8D.4. Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

Solution. Fix an orthonormal basis $e_1, \ldots, e_n$ of V (6.35). Computing both traces from it by 8.55,

$$\operatorname{tr} T^* = \sum_{k=1}^{n} \langle T^* e_k, e_k\rangle = \sum_{k=1}^{n} \overline{\langle T e_k, e_k\rangle} = \overline{\operatorname{tr} T},$$

the middle equality because conjugate symmetry and the defining property 7.1 of the adjoint give $\langle T^* e_k, e_k\rangle = \overline{\langle e_k, T^* e_k\rangle} = \overline{\langle T e_k, e_k\rangle}$, and the last because conjugation is additive.
Method (2): with $A = \mathcal{M}\big(T, (e_1, \ldots, e_n)\big)$, orthonormality of the basis gives $\mathcal{M}(T^*) = A^*$ by 7.9, and

$(A^*)_{k,k} = \overline{A_{k,k}}$ by the definition 7.7 of conjugate transpose. Hence

$$\operatorname{tr} T^* = \sum_{k=1}^{n} \overline{A_{k,k}} = \overline{\sum_{k=1}^{n} A_{k,k}} = \overline{\operatorname{tr} T},$$

the outer equalities by 8.51 computed with this basis, which is legitimate by 8.50.

Problem 8D.5. Suppose V is an inner product space. Suppose $T \in \mathcal{L}(V)$ is a positive operator and $\operatorname{tr} T = 0$. Prove that $T = 0$.

Solution. Diagonalize T: positivity and 7.38 (a) $\Rightarrow$ (c) give an orthonormal basis $e_1, \ldots, e_n$ of V with

$$\mathcal{M}\big(T, (e_1, \ldots, e_n)\big) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \qquad \lambda_k \geq 0.$$

Since the trace may be computed from the matrix with respect to any basis (8.50, 8.51), $0 = \operatorname{tr} T = \lambda_1 + \cdots + \lambda_n$, and a sum of nonnegative reals vanishes only if every term does. So the matrix is the zero matrix, giving $Te_k = 0$ for each k and hence, by linearity on a basis, $T = 0$.
Method (2): by 7.38 (a) $\Rightarrow$ (f) write $T = R^*R$ for some $R \in \mathcal{L}(V)$. For any orthonormal basis $e_1, \ldots, e_n$ (6.35), the defining property 7.1 of the adjoint applied to R^* together with $(R^*)^* = R$ (7.5(c)) gives $\langle R^*Re_k, e_k\rangle = \|Re_k\|^2$, so by 8.55

$$0 = \operatorname{tr} T = \sum_{k=1}^{n} \langle R^*Re_k, e_k\rangle = \sum_{k=1}^{n} \|Re_k\|^2.$$

Each summand is nonnegative, so $Re_k = 0$ for every k, whence $R = 0$ and $T = R^*R = 0$.

Problem 8D.6. Suppose V is an inner product space and $P, Q \in \mathcal{L}(V)$ are orthogonal projections. Prove that $\operatorname{tr}(PQ) \geq 0$.

Solution. Setting $S = QP$, the trace in question equals $\operatorname{tr}(S^*S) = \sum_k \|QPe_k\|^2 \geq 0$ for any orthonormal basis $e_1, \ldots, e_n$ of V (6.35).
Orthogonal projections are idempotent and self-adjoint. Say $P = P_U$ and $Q = P_W$; then $P^2 = P$ and $Q^2 = Q$ by 6.57(g). For $P^* = P$, write $v = u_1 + x_1$ and $w = u_2 + x_2$ with $u_1, u_2 \in U$ and $x_1, x_2 \in U^\perp$ (6.49), so that $Pv = u_1$ and $Pw = u_2$ (6.55); then $\langle u_1, x_2\rangle = \langle x_1, u_2\rangle = 0$ gives

$$\langle Pv, w\rangle = \langle u_1, u_2\rangle = \langle v, Pw\rangle$$

for all $v, w \in V$, so uniqueness of the adjoint yields $P^* = P$, and likewise $Q^* = Q$.
Hence, by 7.5(d) and by 8.56 applied with $A = PQ$, $B = P$,

$$S^*S = P^*Q^*QP = PQ^2P = PQP,$$

$$\operatorname{tr}(S^*S) = \operatorname{tr}\big((PQ)P\big) = \operatorname{tr}\big(P(PQ)\big) = \operatorname{tr}(P^2Q) = \operatorname{tr}(PQ).$$

Finally, by 8.55 together with 7.1 and $S^{**} = S$ (7.5(c)),

$$\operatorname{tr}(S^*S) = \sum_{k=1}^{n} \langle S^*Se_k, e_k\rangle = \sum_{k=1}^{n} \|Se_k\|^2 \geq 0.$$

Problem 8D.7. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

$$\begin{pmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{pmatrix}.$$

Someone tells you (accurately) that -48 and 24 are eigenvalues of T. Without using a computer or writing anything down, find the third eigenvalue of T.

Solution. The third eigenvalue is 36: add the diagonal entries to get $\operatorname{tr} T = 51 - 40 + 1 = 12$, then subtract the two known eigenvalues, $12 - (-48) - 24 = 36$.
This works because the trace is the sum of the diagonal entries (8.47, 8.51) and, over $\mathbf{C}$, also the sum of the eigenvalues counted with multiplicity (8.52), the multiplicities summing to $\dim \mathbf{C}^3 = 3$ (8.25). Since -48 and 24 each have multiplicity at least 1, there are two cases.
(i) They are the only eigenvalues. Then their multiplicities are $2, 1$ or $1, 2$, giving eigenvalue sums

$$2(-48) + 24 = -72 \qquad \text{or} \qquad (-48) + 2(24) = 0,$$

neither equal to 12; so 8.52 rules this out.
(ii) There is a third eigenvalue λ. Then three distinct eigenvalues have positive multiplicities summing to 3, so each multiplicity is 1, and 8.52 gives $12 = (-48) + 24 + \lambda$, that is, $\lambda = 36$.

Problem 8D.8. Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\operatorname{tr}(ST) = (\operatorname{tr} S)(\operatorname{tr} T)$.

Solution. False. Take $V = \mathbf{F}^2$ and $S = T = I$. Then $\operatorname{tr} I = 2$ by 8.47 and 8.51 (the matrix of I in the standard basis is the identity), so

$$\operatorname{tr}(ST) = \operatorname{tr} I = 2 \neq 4 = 2 \cdot 2 = (\operatorname{tr} S)(\operatorname{tr} T).$$

Problem 8D.9. Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{tr}(ST) = 0$ for all $S \in \mathcal{L}(V)$. Prove that $T = 0$.

Solution. Test the hypothesis against the coordinate operators. Fix a basis $v_1, \ldots, v_n$ of V, let $A = \mathcal{M}(T)$ with respect to it, and for $j, k \in \{1, \ldots, n\}$ let $P_{j,k} \in \mathcal{L}(V)$ send $a_1 v_1 + \cdots + a_n v_n \mapsto a_k v_j$ (linear, since the coefficient functionals are). As $P_{j,k} v_k = v_j$ and $P_{j,k} v_i = 0$ for $i \neq k$, we get $\mathcal{M}(P_{j,k}) = E_{j,k}$, the matrix with 1 in row j, column k and 0 elsewhere. Hence by 3.43,

$$(E_{j,k} A)_{r,s} = \sum_{t=1}^{n} (E_{j,k})_{r,t} A_{t,s} = \begin{cases} A_{k,s} & \text{if } r = j, \\ 0 & \text{if } r \neq j, \end{cases}$$

so the only possibly nonzero diagonal entry of $\mathcal{M}(P_{j,k} T) = E_{j,k} A$ sits in row j, where it equals $A_{k,j}$. Thus by 8.47 and 8.51,

$$0 = \operatorname{tr}(P_{j,k} T) = A_{k,j}$$

for all j, k. Hence $\mathcal{M}(T) = 0$, so $T v_k = 0$ for each k, and therefore $T = 0$.

Problem 8D.10. Prove that the trace is the only linear functional $\tau \colon \mathcal{L}(V) \to \mathbf{F}$ such that

$$\tau(ST) = \tau(TS)$$

for all $S, T \in \mathcal{L}(V)$ and $\tau(I) = \dim V$.
Hint: Suppose that $v_1, \ldots, v_n$ is a basis of V. For $j, k \in \{1, \ldots, n\}$, define $P_{j,k} \in \mathcal{L}(V)$ by $P_{j,k}(a_1 v_1 +$

$\cdots + a_n v_n) = a_k v_j$. Prove that

$$\tau(P_{j,k}) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Then for $T \in \mathcal{L}(V)$, use the equation $T = \sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} P_{j,k}$ to show that $\tau(T) = \operatorname{tr} T$.

Solution. The trace itself qualifies: it is a linear functional with $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ by 8.56, and $\operatorname{tr} I = n = \dim V$ by 8.47 and 8.51, where $n = \dim V$. For uniqueness, suppose τ is linear with $\tau(ST) = \tau(TS)$ for all S, T and $\tau(I) = n$. Fix a basis $v_1, \ldots, v_n$ of V and let $P_{j,k} \in \mathcal{L}(V)$ send $a_1 v_1 + \cdots + a_n v_n \mapsto a_k v_j$; equivalently $P_{j,k} v_k = v_j$ and $P_{j,k} v_i = 0$ for $i \neq k$. Evaluating on each v_i gives the multiplication rule (Check!)

$$P_{j,k} P_{l,m} = \begin{cases} P_{j,m} & \text{if } k = l, \\ 0 & \text{if } k \neq l. \end{cases}$$

(i) If $j \neq k$, then $P_{j,j} P_{j,k} = P_{j,k}$ while $P_{j,k} P_{j,j} = 0$, so

$$\tau(P_{j,k}) = \tau(P_{j,j} P_{j,k}) = \tau(P_{j,k} P_{j,j}) = \tau(0) = 0.$$

(ii) For any j, k, the rule gives $P_{j,k} P_{k,j} = P_{j,j}$ and $P_{k,j} P_{j,k} = P_{k,k}$, whence

$$\tau(P_{j,j}) = \tau(P_{j,k} P_{k,j}) = \tau(P_{k,j} P_{j,k}) = \tau(P_{k,k}) =: c.$$

Since $\sum_{j=1}^{n} P_{j,j} = I$ (both sides fix every $v \in V$), linearity yields $n = \tau(I) = nc$; as $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$ and $n \geq 1$, we get $c = 1$.
Now let $T \in \mathcal{L}(V)$ and $A = \mathcal{M}(T, (v_1, \ldots, v_n))$. Both sides of $T = \sum_k \sum_j A_{j,k} P_{j,k}$ send v_i to $\sum_j A_{j,i} v_j$, so they are equal; applying τ and using (i) and (ii),

$$\tau(T) = \sum_{k=1}^{n} \sum_{j=1}^{n} A_{j,k}\, \tau(P_{j,k}) = \sum_{j=1}^{n} A_{j,j} = \operatorname{tr} T$$

by 8.51. Hence $\tau = \operatorname{tr}$.

Problem 8D.11. Suppose V and W are inner product spaces and $T \in \mathcal{L}(V, W)$. Prove that if $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W, then

$$\operatorname{tr}(T^* T) = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle T e_k, f_j \rangle|^2.$$

The numbers $\langle T e_k, f_j \rangle$ are the entries of the matrix of T with respect to the orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_m$. These numbers depend on the bases, but $\operatorname{tr}(T^* T)$ does not depend on a choice of bases. Thus this exercise shows that the sum of the squares of the absolute values of the matrix entries does not depend on which orthonormal bases are used.

Solution. Apply 8.55 to $T^* T \in \mathcal{L}(V)$ and the orthonormal basis $e_1, \ldots, e_n$, then Parseval's identity 6.30(b) to each $T e_k \in W$ and the orthonormal basis $f_1, \ldots, f_m$:

$$\begin{aligned} \operatorname{tr}(T^* T) &= \sum_{k=1}^{n} \langle T^*(T e_k), e_k \rangle = \sum_{k=1}^{n} \overline{\langle e_k, T^*(T e_k) \rangle} \\ &= \sum_{k=1}^{n} \overline{\langle T e_k, T e_k \rangle} = \sum_{k=1}^{n} \|T e_k\|^2 = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle T e_k, f_j \rangle|^2, \end{aligned}$$

the third equality by the definition 7.1 of the adjoint and the fourth because $\|T e_k\|^2$ is real.

Problem 8D.12. Suppose V and W are finite-dimensional inner product spaces.
(a) Prove that $\langle S, T\rangle = \operatorname{tr}(T^*S)$ defines an inner product on $\mathcal{L}(V, W)$.
(b) Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W. Show that the inner product on $\mathcal{L}(V, W)$ from (a) is the same as the standard inner product on $\mathbf{F}^{mn}$, where we identify each element of $\mathcal{L}(V, W)$ with its matrix (with respect to the bases just mentioned) and then with an element of $\mathbf{F}^{mn}$.
Caution: The norm of a linear map $T \in \mathcal{L}(V, W)$ as defined by 7.86 is not the same as the norm that comes from the inner product in (a) above. Unless explicitly stated otherwise, always assume that $\|T\|$ refers to the norm as defined by 7.86. The norm that comes from the inner product in (a) is called the Frobenius norm or the Hilbert–Schmidt norm.

Solution. Throughout fix orthonormal bases $e_1, \ldots, e_n$ of V and $f_1, \ldots, f_m$ of W (available by 6.35), and note $T^*S \in \mathcal{L}(V)$ by 7.4.
(a) The four axioms of 6.2 all hold.
(i) Positivity and definiteness: by Exercise 8D.11,

$$\langle S, S\rangle = \operatorname{tr}(S^*S) = \sum_{k=1}^{n} \|Se_k\|^2 \geq 0,$$

which vanishes exactly when $Se_k = 0$ for every k, i.e. when $S = 0$.
(ii) Additivity and homogeneity in the first slot: since $T^*(S_1 + \lambda S_2) = T^*S_1 + \lambda T^*S_2$, linearity of the trace (8.56) gives

$$\langle S_1 + \lambda S_2, T\rangle = \langle S_1, T\rangle + \lambda\langle S_2, T\rangle.$$

(iii) Conjugate symmetry: $(T^*S)^* = S^*T$ by 7.5, so by Exercise 8D.4,

$$\langle T, S\rangle = \operatorname{tr}\big((T^*S)^*\big) = \overline{\operatorname{tr}(T^*S)} = \overline{\langle S, T\rangle}.$$

(b) By 6.30(a) the entries of $A = \mathcal{M}(T, (e_1, \ldots, e_n), (f_1, \ldots, f_m))$ and of the matrix B of S are $A_{j,k} = \langle Te_k, f_j\rangle$ and $B_{j,k} = \langle Se_k, f_j\rangle$. Now 8.55, the definition 7.1 of the adjoint, and 6.30(c) applied in W give

$$\begin{aligned}\langle S, T\rangle &= \sum_{k=1}^{n} \langle T^*(Se_k), e_k\rangle = \sum_{k=1}^{n} \langle Se_k, Te_k\rangle \\ &= \sum_{k=1}^{n} \sum_{j=1}^{m} B_{j,k}\overline{A_{j,k}},\end{aligned}$$

which under any listing of the mn pairs (j, k) is the standard inner product of the vectors in $\mathbf{F}^{mn}$ corresponding to S and T. Since $\mathcal{M}$ is an isomorphism of $\mathcal{L}(V, W)$ onto $\mathbf{F}^{m,n}$ (3.71), the two inner products agree under this identification.

Problem 8D.13. Find $S, T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ such that $ST - TS = I$.
Hint: Make an appropriate modification of the operators in Example 3.9.
This exercise shows that additional hypotheses are needed on S and T to extend 8.57 to the setting of infinite-dimensional vector spaces.

Solution. Take $S =$ differentiation and $T =$ multiplication by x:

$$(Sp)(x) = p'(x), \qquad (Tp)(x) = x\,p(x).$$

Both map $\mathcal{P}(\mathbf{F})$ into itself and are linear (Check!; over $\mathbf{C}$ read p' as the formal derivative). By the product rule,

$$\big((ST - TS)p\big)(x) = \big(x\,p(x)\big)' - x\,p'(x) = p(x) + x\,p'(x) - x\,p'(x) = p(x)$$

for every $p \in \mathcal{P}(\mathbf{F})$, so $ST - TS = I$.

9 Multilinear Algebra and Determinants

Contents

9.1 Exercises 9A

Problem 9A.1. Prove that if β is a bilinear form on $\mathbf{F}$, then there exists $c \in \mathbf{F}$ such that

$$\beta(x, y) = c\,x\,y$$

for all $x, y \in \mathbf{F}$.

Solution. Take $c = \beta(1, 1)$. Writing $x = x \cdot 1$ and $y = y \cdot 1$ and using homogeneity of β in the first slot and then in the second,

$$\beta(x, y) = \beta(x \cdot 1,\ y \cdot 1) = x\,\beta(1,\ y \cdot 1) = x\,y\,\beta(1, 1) = c\,x\,y$$

for all $x, y \in \mathbf{F}$.

Problem 9A.2. Let $n = \dim V$. Suppose β is a bilinear form on V. Prove that there exist $\varphi_1, \dots, \varphi_n, \tau_1, \dots, \tau_n \in V'$ such that

$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \dots + \varphi_n(u) \cdot \tau_n(v)$$

for all $u, v \in V$.
[This exercise shows that if $n = \dim V$, then every bilinear form on V is of the form given by the last bullet point of Example 9.2.]

Solution. Take $\varphi_1, \dots, \varphi_n$ to be the dual basis (3.112) of a basis $e_1, \dots, e_n$ of V, and set $\tau_k(v) = \beta(e_k, v)$; each τ_k lies in V' because β is linear in its second slot. Since $u = \varphi_1(u)e_1 + \dots + \varphi_n(u)e_n$ by 3.114, linearity of β in its first slot gives

$$\begin{aligned}\beta(u, v) &= \beta\big(\varphi_1(u)\,e_1 + \dots + \varphi_n(u)\,e_n,\ v\big)\\ &= \varphi_1(u)\,\beta(e_1, v) + \dots + \varphi_n(u)\,\beta(e_n, v)\\ &= \varphi_1(u) \cdot \tau_1(v) + \dots + \varphi_n(u) \cdot \tau_n(v)\end{aligned}$$

for all $u, v \in V$.

Problem 9A.3. Suppose $\beta\colon V \times V \to \mathbf{F}$ is a bilinear form on V and also is a linear functional on $V \times V$. Prove that $\beta = 0$.

Solution. Split $(u, v) = (u, 0) + (0, v)$ in $V \times V$. Bilinearity makes $w \mapsto \beta(u, w)$ and $w \mapsto \beta(w, u)$ linear maps, and linear maps send 0 to 0, so $\beta(u, 0) = \beta(0, v) = 0$. Additivity of β on $V \times V$ then gives

$$\beta(u, v) = \beta\big((u, 0) + (0, v)\big) = \beta(u, 0) + \beta(0, v) = 0$$

for all $u, v \in V$. Hence $\beta = 0$.

Problem 9A.4. Suppose V is a real inner product space and β is a bilinear form on V. Show that there exists a unique operator $T \in \mathcal{L}(V)$ such that

$$\beta(u, v) = \langle u, Tv \rangle$$

for all $u, v \in V$.
[This exercise states that if V is a real inner product space, then every bilinear form on V is of the form given by the third bullet point in 9.2.]

Solution. Define Tv by Riesz representation. For each $v \in V$ the map $u \mapsto \beta(u, v)$ is a linear functional on the finite-dimensional space V, so by 6.42 there is a unique vector $Tv \in V$ with

$$\beta(u, v) = \langle u, Tv\rangle \quad \text{for all } u \in V,$$

which defines a function $T\colon V \to V$. Throughout we use that $\langle u, w\rangle = 0$ for all $u \in V$ forces $w = 0$ (take $u = w$ and apply definiteness in 6.2).
T is linear: for $v_1, v_2 \in V$, $\lambda \in \mathbf{R}$ and every $u \in V$, additivity and homogeneity of β in its second slot give

$$\langle u, T(v_1 + v_2)\rangle = \beta(u, v_1) + \beta(u, v_2) = \langle u, Tv_1 + Tv_2\rangle,$$
$$\langle u, T(\lambda v)\rangle = \lambda\, \beta(u, v) = \lambda\langle u, Tv\rangle = \langle u, \lambda Tv\rangle,$$

the last step using that a real inner product is homogeneous (not conjugate-homogeneous) in its second slot; hence $T(v_1 + v_2) = Tv_1 + Tv_2$ and $T(\lambda v) = \lambda Tv$.
Uniqueness: if S, T both represent β, then $\langle u, Sv - Tv\rangle = 0$ for all $u, v \in V$, so $S = T$.

Problem 9A.5. Suppose β is a bilinear form on a real inner product space V and T is the unique operator on V such that $\beta(u, v) = \langle u, Tv\rangle$ for all $u, v \in V$ (see Exercise 4). Show that β is an inner product on V if and only if T is an invertible positive operator on V.

Solution. Since β is bilinear, additivity and homogeneity in the first slot are automatic, so by 6.2 the question is whether β is symmetric, positive, and definite.
(i) Suppose T is invertible and positive, so $T^* = T$ and $\langle Tv, v\rangle \geq 0$ (7.34). Symmetry and positivity follow at once from symmetry of the real inner product:

$$\beta(u, v) = \langle u, Tv\rangle = \langle Tu, v\rangle = \langle v, Tu\rangle = \beta(v, u), \qquad \beta(v, v) = \langle Tv, v\rangle \geq 0.$$

For definiteness, take a positive R with $R^2 = T$ (7.38, (a) $\Rightarrow$ (d)); R is self-adjoint, so $\beta(v, v) = 0$ gives

$$0 = \langle R^2 v, v\rangle = \langle Rv, R^* v\rangle = \|Rv\|^2,$$

whence $Rv = 0$, so $Tv = R(Rv) = 0$ and $v = 0$ by injectivity of T. Thus β is an inner product.
(ii) Suppose β is an inner product. Symmetry of β and of $\langle\cdot, \cdot\rangle$ gives

$$\langle u, Tv\rangle = \beta(u, v) = \beta(v, u) = \langle v, Tu\rangle = \langle Tu, v\rangle = \langle u, T^* v\rangle$$

for all u, v; taking $u = Tv - T^* v$ yields $T^* = T$. Positivity of β gives $\langle Tv, v\rangle = \beta(v, v) \geq 0$, so T is a positive operator by 7.34. Finally, $Tv = 0$ forces $\beta(v, v) = \langle v, 0\rangle = 0$ and hence $v = 0$, so T is injective and therefore invertible (3.65, V being finite-dimensional).

Problem 9A.6. Prove or give a counterexample: If ρ is a symmetric bilinear form on V, then

$$\{v \in V : \rho(v, v) = 0\}$$

is a subspace of V.

Solution. False. On $V = \mathbf{R}^2$ take

$$\rho\big((x_1, x_2), (y_1, y_2)\big) = x_1 y_1 - x_2 y_2, \qquad \mathcal{M}(\rho) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

a bilinear form (each slot is a combination of coordinate functionals) whose matrix with respect to the standard basis is symmetric, so ρ is symmetric by 9.12. Then $(1, 1)$ and $(1, -1)$ both lie in $\{v : \rho(v, v) = 0\}$, since $1 - 1 = 0$ in each case, but their sum does not:

$$\rho\big((2, 0), (2, 0)\big) = 4 - 0 = 4 \neq 0.$$

So the set is not closed under addition, hence is not a subspace.

Problem 9A.7. Explain why the proof of 9.13 (diagonalization of a symmetric bilinear form by an orthonormal basis on a real inner product space) fails if the hypothesis that $\mathbf{F} = \mathbf{R}$ is dropped.

Solution. Three steps of the proof of 9.13 use $\mathbf{F} = \mathbf{R}$, and each breaks over $\mathbf{C}$. That proof takes $B = \mathcal{M}(\rho, (f_1, \ldots, f_n))$ for an orthonormal basis $f_1, \ldots, f_n$ (symmetric by 9.12), lets T be the operator with that matrix, calls T self-adjoint, diagonalizes T by the real spectral theorem, and concludes with $\mathcal{M}(\rho, (e_1, \ldots, e_n)) = C^t BC = C^{-1}BC$.
(i) "B symmetric $\Rightarrow$ T self-adjoint" fails: by 7.9 the matrix of T^* with respect to an orthonormal basis is the conjugate transpose, so T is self-adjoint iff B is Hermitian, which for symmetric B means all entries real. On $V = \mathbf{C}$ with $f_1 = 1$ and $\rho(u, w) = i\,uw$, the matrix (i) is symmetric but T is multiplication by i, with adjoint multiplication by $-i$.
(ii) The real spectral theorem 7.29 requires $\mathbf{F} = \mathbf{R}$, and its complex substitute 7.31 requires T normal, which a complex symmetric B need not give. On $\mathbf{C}^2$ with the standard basis and

$$B = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}, \quad B^*B = \begin{pmatrix} 2 & 2i \\ -2i & 2 \end{pmatrix} \neq \begin{pmatrix} 2 & -2i \\ 2i & 2 \end{pmatrix} = BB^*,$$

so T is not normal and neither spectral theorem applies.
(iii) Even granted a diagonalizing orthonormal basis, the last step collapses: a change of basis transforms the matrix of a bilinear form to C^tBC (9.7) but that of an operator to $C^{-1}BC$ (3.84). These agree over $\mathbf{R}$ because a real unitary C has $C^{-1} = C^t$ (7.57); over $\mathbf{C}$ a unitary C has $C^{-1} = \overline{C^t} \neq C^t$ in general, so diagonalizing $C^{-1}BC$ says nothing about C^tBC.

Problem 9A.8. Find formulas for $\dim V^{(2)}_{\mathrm{sym}}$ and $\dim V^{(2)}_{\mathrm{alt}}$ in terms of $\dim V$.

Solution. With $n = \dim V$,

$$\dim V^{(2)}_{\mathrm{sym}} = \frac{n(n+1)}{2}, \qquad \dim V^{(2)}_{\mathrm{alt}} = \frac{n(n-1)}{2}.$$

Fix a basis $e_1, \ldots, e_n$ of V; by 9.5 the map $\beta \mapsto \mathcal{M}(\beta)$ is an isomorphism of $V^{(2)}$ onto $\mathbf{F}^{n,n}$, and by 9.12 ((a) $\Leftrightarrow$ (c) there) it carries $V^{(2)}_{\mathrm{sym}}$ onto

$$\mathcal{S} = \{A \in \mathbf{F}^{n,n} : A_{j,k} = A_{k,j} \text{ for all } j, k\}.$$

Let $E_{j,k}$ be the matrix with 1 in row j, column k and 0 elsewhere. The matrices $E_{j,j}$ for $j \in \{1, \ldots, n\}$ together with $E_{j,k} + E_{k,j}$ for $j < k$ all lie in $\mathcal{S}$ and form a basis of it: they span, since for $A \in \mathcal{S}$

$$A = \sum_{j=1}^{n} A_{j,j}E_{j,j} + \sum_{j<k} A_{j,k}(E_{j,k} + E_{k,j})$$

entry by entry (using $A_{k,j} = A_{j,k}$), and they are independent because the entry in row j, column k reads off each coefficient (Check!). Counting the n diagonal matrices and the $\binom{n}{2}$ pairs,

$$\dim V^{(2)}_{\mathrm{sym}} = \dim \mathcal{S} = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

Finally $V^{(2)} = V^{(2)}_{\mathrm{sym}} \oplus V^{(2)}_{\mathrm{alt}}$ by 9.17 and $\dim V^{(2)} = n^2$ by 9.5, so

$$\dim V^{(2)}_{\mathrm{alt}} = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

Problem 9A.9. Suppose that n is a positive integer and $V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}$. Define $\alpha : V \times V \to \mathbf{R}$ by

$$\alpha(p,q) = \int_0^1 pq'.$$

Show that α is an alternating bilinear form on V.

Solution. Bilinearity comes from linearity of the integral and of differentiation: for $p_1, p_2, q_1, q_2 \in V$ and $\lambda \in \mathbf{R}$,

$$\begin{aligned}\alpha(p_1 + \lambda p_2, q) &= \int_0^1 (p_1 q' + \lambda p_2 q') \\ &= \alpha(p_1, q) + \lambda\, \alpha(p_2, q), \\ \alpha(p, q_1 + \lambda q_2) &= \int_0^1 p\,(q_1' + \lambda q_2') \\ &= \alpha(p, q_1) + \lambda\, \alpha(p, q_2),\end{aligned}$$

so α is a bilinear form on V (9.1). For the alternating condition 9.14, the product rule gives $pp' = \frac{1}{2}(p^2)'$, and p^2 is a polynomial, hence continuously differentiable on $[0,1]$, so the fundamental theorem of calculus yields

$$\alpha(p,p) = \frac{1}{2}\int_0^1 (p^2)' = \frac{p(1)^2 - p(0)^2}{2} = 0$$

for every $p \in V$, the last equality because $p(0) = p(1)$.

Problem 9A.10. Suppose that n is a positive integer and

$$V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1) \text{ and } p'(0) = p'(1)\}.$$

Define $\rho : V \times V \to \mathbf{R}$ by

$$\rho(p,q) = \int_0^1 pq''.$$

Show that ρ is a symmetric bilinear form on V.

Solution. Integration by parts turns ρ into a manifestly symmetric expression. Bilinearity is immediate from linearity of the integral and of $q \mapsto q''$:

$$\begin{aligned}\rho(p_1 + \lambda p_2, q) &= \int_0^1 (p_1 q'' + \lambda p_2 q'') \\ &= \rho(p_1, q) + \lambda\, \rho(p_2, q), \\ \rho(p, q_1 + \lambda q_2) &= \int_0^1 p\,(q_1'' + \lambda q_2'') \\ &= \rho(p, q_1) + \lambda\, \rho(p, q_2),\end{aligned}$$

so ρ is a bilinear form (9.1). For symmetry, pq' is a polynomial, hence continuously differentiable on $[0,1]$, so the product rule and the fundamental theorem of calculus give

$$\int_0^1 (p'q' + pq'') = \int_0^1 (pq')' = p(1)q'(1) - p(0)q'(0) = 0,$$

the last equality because $p(0) = p(1)$ and $q'(0) = q'(1)$. Hence $\rho(p,q) = -\int_0^1 p'q'$ for all $p, q \in V$; the same formula with p and q interchanged (legitimate, as $q(0) = q(1)$ and $p'(0) = p'(1)$ too) gives

$$\rho(q,p) = -\int_0^1 q'p' = -\int_0^1 p'q' = \rho(p,q),$$

which is the symmetry condition 9.9.

9.2 Exercises 9B

Problem 9B.1. Suppose m is a positive integer. Show that $\dim V^{(m)} = (\dim V)^m$.

Solution. The n^m products of dual-basis functionals form a basis of $V^{(m)}$, where $n = \dim V$. Fix a basis $e_1, \dots, e_n$ of V with dual basis $\varphi_1, \dots, \varphi_n$, so $v = \sum_j \varphi_j(v)e_j$ by 3.114, and for each list $(j_1, \dots, j_m)$ in $\{1, \dots, n\}$ set

$$\beta_{j_1,\dots,j_m}(v_1, \dots, v_m) = \varphi_{j_1}(v_1) \cdots \varphi_{j_m}(v_m).$$

Each lies in $V^{(m)}$, since freeing the k-th slot leaves $v \mapsto c\,\varphi_{j_k}(v)$ with $c = \prod_{p \neq k} \varphi_{j_p}(u_p)$ independent of v.

Spanning: for $\beta \in V^{(m)}$, expanding each v_k in the basis and using linearity in each slot in turn gives

$$\begin{aligned}\beta(v_1, \dots, v_m) &= \sum_{j_1=1}^{n} \cdots \sum_{j_m=1}^{n} \varphi_{j_1}(v_1) \cdots \varphi_{j_m}(v_m)\, \beta(e_{j_1}, \dots, e_{j_m}) \\ &= \sum_{j_1=1}^{n} \cdots \sum_{j_m=1}^{n} \beta(e_{j_1}, \dots, e_{j_m})\, \beta_{j_1,\dots,j_m}(v_1, \dots, v_m),\end{aligned}$$

so $\beta = \sum_{j_1} \cdots \sum_{j_m} \beta(e_{j_1}, \dots, e_{j_m})\, \beta_{j_1,\dots,j_m}$.

Independence: evaluating a vanishing combination $\sum c_{j_1,\dots,j_m}\beta_{j_1,\dots,j_m} = 0$ at $(e_{k_1}, \dots, e_{k_m})$ kills every term except the one with $j_p = k_p$ for all p, giving $c_{k_1,\dots,k_m} = 0$.

There are n^m such index lists, so $\dim V^{(m)} = n^m = (\dim V)^m$.

Problem 9B.2. Suppose $n \geq 3$ and $\alpha \colon \mathbf{F}^n \times \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ is defined by

$$\begin{aligned}&\alpha\big((x_1, \dots, x_n), (y_1, \dots, y_n), (z_1, \dots, z_n)\big) \\ &\qquad = x_1y_2z_3 - x_2y_1z_3 - x_3y_2z_1 - x_1y_3z_2 + x_3y_1z_2 + x_2y_3z_1.\end{aligned}$$

Show that α is an alternating 3-linear form on $\mathbf{F}^n$.

Solution. The six terms are exactly the six permutation terms of perm 3, each weighted by its sign. Writing φ_j for the j-th coordinate functional on $\mathbf{F}^n$ and computing signs from 9.32,

$$\begin{aligned}&\operatorname{sign}(1,2,3) = 1, && \operatorname{sign}(2,1,3) = -1, \\ &\operatorname{sign}(3,2,1) = -1, && \operatorname{sign}(1,3,2) = -1, \\ &\operatorname{sign}(3,1,2) = 1, && \operatorname{sign}(2,3,1) = 1,\end{aligned}$$

so matching term by term with the defining formula gives

$$\alpha(u, v, w) = \sum_{(j_1,j_2,j_3) \in \operatorname{perm} 3} \big(\operatorname{sign}(j_1, j_2, j_3)\big)\, \varphi_{j_1}(u)\, \varphi_{j_2}(v)\, \varphi_{j_3}(w).$$

3-linearity: freeing one slot leaves a linear combination of $\varphi_1, \varphi_2, \varphi_3$ whose coefficients involve only the other two (fixed) vectors, hence a linear functional on $\mathbf{F}^n$.

Alternating: by 9.27 it suffices that $\alpha(u, v, w) = 0$ when two arguments coincide. If $u = v$, the map $(j_1, j_2, j_3) \mapsto (j_2, j_1, j_3)$ is a fixed-point-free involution of perm 3, so it splits perm 3 into three pairs; it reverses sign by 9.34 while leaving $\varphi_{j_1}(u)\varphi_{j_2}(v)\varphi_{j_3}(w)$ unchanged (since $u = v$), so the paired terms cancel and the sum is 0. For $v = w$ use $(j_1, j_2, j_3) \mapsto (j_1, j_3, j_2)$; for $u = w$ use $(j_1, j_2, j_3) \mapsto (j_3, j_2, j_1)$.

Problem 9B.3. Suppose m is a positive integer and α is an m-linear form on V such that $\alpha(v_1, \dots, v_m) = 0$ whenever $v_1, \dots, v_m$ is a list of vectors in V with $v_j = v_{j+1}$ for some $j \in \{1, \dots, m-1\}$. Prove that α is an alternating m-linear form on V.

Solution. Swapping adjacent slots reverses the sign of α, and that reduces equal entries in any two slots to the adjacent case. (For $m = 1$ the condition in 9.27 is vacuous, so assume $m \geq 2$.)
(i) Fix $j \in \{1, \dots, m-1\}$ and all entries outside slots $j, j+1$, and write $\alpha(\dots, a, b, \dots)$ for the value with a in slot j and b in slot $j+1$. Putting $u + w$ in both slots and expanding by linearity in each,

$$0 = \alpha(\dots, u, u, \dots) + \alpha(\dots, u, w, \dots) + \alpha(\dots, w, u, \dots) + \alpha(\dots, w, w, \dots),$$

where the outer terms vanish by hypothesis; hence $\alpha(\dots, w, u, \dots) = -\alpha(\dots, u, w, \dots)$.
(ii) Suppose $v_j = v_k$ with $j < k$. If $k = j+1$, the hypothesis gives $\alpha(v_1, \dots, v_m) = 0$ directly. If $k > j+1$, perform the $k-j-1$ adjacent swaps of slots $(k-1, k)$, then $(k-2, k-1)$, down to $(j+1, j+2)$; these never touch slot j and move v_k into slot $j+1$, producing the list

$$v_1, \dots, v_{j-1}, v_j, v_k, v_{j+1}, \dots, v_{k-1}, v_{k+1}, \dots, v_m,$$

whose slots j and $j+1$ hold the equal vectors v_j, v_k. So α of it is 0, and by (i)

$$\alpha(v_1, \dots, v_m) = (-1)^{k-j-1} \cdot 0 = 0.$$

By 9.27, α is alternating.

Problem 9B.4. Prove or give a counterexample: If $\alpha \in V_{\text{alt}}^{(4)}$, then

$$\{(v_1, v_2, v_3, v_4) \in V^4 : \alpha(v_1, v_2, v_3, v_4) = 0\}$$

is a subspace of V^4.

Solution. False. Let $V = \mathbf{F}^4$ with standard basis e_1, e_2, e_3, e_4, and let $\alpha \in V_{\text{alt}}^{(4)}$ be the form built in 9.38 from that basis, so $\alpha(e_1, e_2, e_3, e_4) = 1$. Set

$$E = \{(v_1, v_2, v_3, v_4) \in V^4 : \alpha(v_1, v_2, v_3, v_4) = 0\}.$$

Then $(e_1, e_2, e_3, e_3) \in E$, since α is alternating (9.27) and two entries coincide, and $(0, 0, 0, e_4) \in E$, since α is linear in its first slot. But their coordinatewise sum is not in E: by linearity in the fourth slot,

$$\alpha(e_1, e_2, e_3, e_3 + e_4) = \alpha(e_1, e_2, e_3, e_3) + \alpha(e_1, e_2, e_3, e_4) = 0 + 1 = 1 \neq 0.$$

So E is not closed under addition, hence is not a subspace of V^4.

Problem 9B.5. Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, \dots, v_m) = \sum_{(j_1, \dots, j_m) \in \operatorname{perm} m} \bigl(\operatorname{sign}(j_1, \dots, j_m)\bigr) \beta(v_{j_1}, \dots, v_{j_m})$$

for $v_1, \dots, v_m \in V$. Explain why $\alpha \in V_{\text{alt}}^{(m)}$.

Solution. The sum cancels in pairs under a sign-reversing involution of $\operatorname{perm} m$. Identify $(j_1, \dots, j_m) \in \operatorname{perm} m$ with the bijection j of $\{1, \dots, m\}$ given by $j(p) = j_p$.
Suppose $v_r = v_s$ with $r \neq s$ (for $m = 1$ the condition 9.27 is vacuous). Given $(j_1, \dots, j_m)$, let p, q be the indices with $j_p = r$ and $j_q = s$, which are distinct, and let $\Phi(j_1, \dots, j_m) = (j'_1, \dots, j'_m)$ be obtained by interchanging the entries in slots p and q; it again lies in $\operatorname{perm} m$. Then:
(i) Φ is a fixed-point-free involution, since $j'_p = s \neq r = j_p$ and applying Φ again swaps back those same two slots; so Φ partitions $\operatorname{perm} m$ into two-element subsets.
(ii) Φ reverses sign, by 9.34 applied to a swap of two entries.

(iii) Φ leaves β unchanged: $v_{j'_p} = v_s = v_r = v_{j_p}$ and $v_{j'_q} = v_r = v_s = v_{j_q}$, with $v_{j'_i} = v_{j_i}$ otherwise, so the two argument lists are identical.
Hence each pair $\{(j_1, \ldots, j_m), \Phi(j_1, \ldots, j_m)\}$ contributes

$$\big(\operatorname{sign}(j_1, \ldots, j_m) - \operatorname{sign}(j_1, \ldots, j_m)\big)\beta(v_{j_1}, \ldots, v_{j_m}) = 0$$

to $\alpha(v_1, \ldots, v_m)$, so $\alpha(v_1, \ldots, v_m) = 0$ and 9.27 gives $\alpha \in V^{(m)}_{\mathrm{alt}}$.

Problem 9B.6. Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, \ldots, v_m) = \sum_{(j_1,\ldots,j_m)\in \operatorname{perm} m} \beta(v_{j_1}, \ldots, v_{j_m})$$

for $v_1, \ldots, v_m \in V$. Explain why

$$\alpha(v_{k_1}, \ldots, v_{k_m}) = \alpha(v_1, \ldots, v_m)$$

for all $v_1, \ldots, v_m \in V$ and all $(k_1, \ldots, k_m) \in \operatorname{perm} m$.

Solution. Rearranging the inputs of α only reindexes the sum defining it. Identify each $(j_1, \ldots, j_m) \in \operatorname{perm} m$ with the bijection j of $\{1, \ldots, m\}$ given by $j(p) = j_p$. Putting $w_p = v_{k_p}$, the definition of α gives

$$\begin{aligned}\alpha(v_{k_1}, \ldots, v_{k_m}) &= \sum_{(j_1,\ldots,j_m)\in \operatorname{perm} m} \beta(w_{j_1}, \ldots, w_{j_m})\\ &= \sum_{(j_1,\ldots,j_m)\in \operatorname{perm} m} \beta(v_{k_{j_1}}, \ldots, v_{k_{j_m}}).\end{aligned}$$

Now $\Psi(j_1, \ldots, j_m) = (k_{j_1}, \ldots, k_{j_m})$ is the list of the composition $k \circ j$, again a bijection of $\{1, \ldots, m\}$, so Ψ maps $\operatorname{perm} m$ into itself; and Ψ is injective, since $k_{j_p} = k_{j'_p}$ for all p forces $j = j'$ by injectivity of k, hence bijective on the finite set $\operatorname{perm} m$. Reindexing the finite sum by Ψ therefore gives

$$\alpha(v_{k_1}, \ldots, v_{k_m}) = \sum_{(i_1,\ldots,i_m)\in \operatorname{perm} m} \beta(v_{i_1}, \ldots, v_{i_m}) = \alpha(v_1, \ldots, v_m).$$

Problem 9B.7. Give an example of a nonzero alternating 2-linear form α on $\mathbf{R}^3$ and a linearly independent list v_1, v_2 in $\mathbf{R}^3$ such that $\alpha(v_1, v_2) = 0$.
This exercise shows that 9.39 can fail if the hypothesis that $n = \dim V$ is deleted.

Solution. Take

$$\alpha\big((x_1, x_2, x_3), (y_1, y_2, y_3)\big) = x_1y_2 - x_2y_1, \qquad v_1 = (1, 0, 0), \quad v_2 = (0, 0, 1).$$

Each slot of α is a combination of the first two coordinate functionals with coefficients independent of that slot, so α is a 2-linear form (9.25); it is alternating since $\alpha(u, u) = x_1x_2 - x_2x_1 = 0$ (9.27); and it is nonzero since $\alpha\big((1, 0, 0), (0, 1, 0)\big) = 1$. The list v_1, v_2 is linearly independent ($av_1 + bv_2 = (a, 0, b) = 0$ forces $a = b = 0$), yet

$$\alpha(v_1, v_2) = 1 \cdot 0 - 0 \cdot 0 = 0.$$

9.3 Exercises 9C

Problem 9C.1. Prove or give a counterexample: $S, T \in \mathcal{L}(V) \implies \det(S + T) = \det S + \det T$.

Solution. False. Take $V = \mathbf{F}^2$ and $S = T = I$. The two bullet points of 9.42 give $\det(\lambda I) = \lambda^{\dim V}$ and

$\det I = 1$, so

$$\det(S+T) = \det(2I) = 2^2 = 4 \neq 2 = 1 + 1 = \det S + \det T.$$

Problem 9C.2. Suppose the first column of a square matrix A consists of all zeros except possibly the first entry $A_{1,1}$. Let B be the matrix obtained from A by deleting the first row and the first column of A. Show that $\det A = A_{1,1} \det B$.

Solution. Only the permutations fixing 1 survive in 9.46, and those are $\operatorname{perm}(n-1)$ relabelled. Let $n \geq 2$ be the size of A (so that B is a matrix in the sense of 9.43); the hypothesis is $A_{j,1} = 0$ for $j \geq 2$, and $B_{r,s} = A_{r+1,\,s+1}$. Since a permutation with $j_1 \neq 1$ contributes a factor $A_{j_1,1} = 0$, formula 9.46 collapses to

$$\det A = \sum_{\substack{(j_1,\dots,j_n)\in \operatorname{perm} n \\ j_1 = 1}} \big(\operatorname{sign}(j_1,\dots,j_n)\big) A_{1,1} A_{j_2,2} \cdots A_{j_n,n}.$$

The map $(1, j_2, \dots, j_n) \mapsto (k_1, \dots, k_{n-1}) = (j_2 - 1, \dots, j_n - 1)$ is a bijection of this index set onto $\operatorname{perm}(n-1)$, with inverse $j_{r+1} = k_r + 1$. It preserves signs: by 9.32 the sign is $(-1)^N$ with N the number of pairs $k < \ell$ appearing out of order, no such pair involves the leading 1, and subtracting 1 from every remaining entry is an order-preserving relabelling. It also preserves products, since

$$A_{j_{r+1},\,r+1} = A_{k_r+1,\,r+1} = B_{k_r,r} \qquad \text{for } r \in \{1,\dots,n-1\}.$$

Reindexing accordingly and applying 9.46 to B,

$$\det A = A_{1,1} \sum_{(k_1,\dots,k_{n-1})\in \operatorname{perm}(n-1)} \big(\operatorname{sign}(k_1,\dots,k_{n-1})\big) B_{k_1,1} \cdots B_{k_{n-1},\,n-1} = A_{1,1} \det B.$$

Problem 9C.3. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $\det(I+T) = 1$.

Solution. By 8.18 there is a basis of V with respect to which $N = \mathcal{M}(T)$ is upper triangular with all diagonal entries 0; this holds over $\mathbf{R}$ as well as over $\mathbf{C}$, since the minimal polynomial of a nilpotent operator is z^m, so 5.44 applies over either field and 5.41 makes the diagonal entries 0. With respect to that basis $\mathcal{M}(I+T) = I + N$ is upper triangular with every diagonal entry equal to 1, so 9.53 and 9.48 give

$$\det(I+T) = \det(I+N) = 1 \cdots 1 = 1.$$

Method (2) (for $\mathbf{F} = \mathbf{C}$): if $(I+T)v = \lambda v$ with $v \neq 0$, then $Tv = (\lambda - 1)v$, so $\lambda - 1 = 0$ by 8.17(a). Thus 1 is the only eigenvalue of $I+T$, of multiplicity $\dim V$ by 8.25, and 9.55 gives $\det(I+T) = 1^{\dim V} = 1$.

Problem 9C.4. Suppose V is an inner product space and $S \in \mathcal{L}(V)$. Prove that S is unitary if and only if $|\det S| = \|S\| = 1$.

Solution. Let $n = \dim V$ and let $s_1 \geq \cdots \geq s_n$ be the singular values of S (7.65), so that

$$\|S\| = s_1 \qquad \text{and} \qquad |\det S| = s_1 \cdots s_n$$

by 7.85 with the definition 7.86, and by 9.60.
(i) If S is unitary, then S is an isometry (7.51), so $\|Sv\| = \|v\|$ for every v; since $V \neq \{0\}$ there is a vector of norm 1, so the maximum defining $\|S\|$ equals 1. Also $|\det S| = 1$ by 9.58.
(ii) Conversely, suppose $|\det S| = \|S\| = 1$. Then $s_1 = 1$, so $0 \leq s_k \leq 1$ for every k; were some $s_k < 1$, then $s_1 \cdots s_n \leq s_k < 1$, contradicting $s_1 \cdots s_n = |\det S| = 1$. Hence $s_1 = \cdots = s_n = 1$, and the singular value decomposition 7.70 supplies orthonormal lists $e_1, \dots, e_n$ and $f_1, \dots, f_n$ in V, each a basis (6.25 and 2.38), with

$$Sv = \langle v, e_1 \rangle f_1 + \cdots + \langle v, e_n \rangle f_n$$

for every $v \in V$. Then 6.24 and Parseval's identity 6.30(b) give

$$\|Sv\|^2 = \sum_{k=1}^{n} \big|\langle v, e_k\rangle\big|^2 = \|v\|^2,$$

so S is an isometry, hence injective, hence invertible (3.65), hence unitary (7.51).

Problem 9C.5. Suppose A is a block upper-triangular matrix

$$A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_k along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1) \cdots (\det A_m).$$

Solution. Say A_k is n_k-by-n_k and $n = n_1 + \cdots + n_m$. Deleting the first n_1 rows and columns of A leaves a block upper-triangular matrix with diagonal blocks $A_2, \ldots, A_m$, and everything below A_1 vanishes; so by induction on m it suffices to treat $m = 2$, say

$$A = \begin{pmatrix} A_1 & C \\ 0 & A_2 \end{pmatrix}$$

with A_1 of size p, A_2 of size q, and $n = p + q$; thus $A_{j,k} = 0$ whenever $j > p \geq k$, while $A_{j,k} = (A_1)_{j,k}$ for $j, k \leq p$ and $A_{j,k} = (A_2)_{j-p,k-p}$ for $j, k > p$.
In the formula 9.46

$$\det A = \sum_{(j_1,\ldots,j_n) \in \operatorname{perm} n} \big(\operatorname{sign}(j_1, \ldots, j_n)\big) A_{j_1,1} \cdots A_{j_n,n}$$

a nonzero term forces $j_k \leq p$ for every $k \leq p$, hence forces $\sigma = (j_1, \ldots, j_p) \in \operatorname{perm} p$ and $\tau = (j_{p+1} - p, \ldots, j_n - p) \in \operatorname{perm} q$; conversely each such pair (σ, τ) arises from exactly one permutation of $(1, \ldots, n)$. For such a permutation the values at most p occupy the first p slots, so no pair straddling p is inverted and the inversion count of 9.32 splits:

$$\operatorname{sign}(\sigma_1, \ldots, \sigma_p, \tau_1 + p, \ldots, \tau_q + p) = (\operatorname{sign} \sigma)(\operatorname{sign} \tau).$$

Since also $A_{\sigma_k,k} = (A_1)_{\sigma_k,k}$ for $k \leq p$ and $A_{\tau_r+p,\,p+r} = (A_2)_{\tau_r,r}$, the sum factors:

$$\begin{aligned} \det A &= \sum_{\sigma \in \operatorname{perm} p} \sum_{\tau \in \operatorname{perm} q} (\operatorname{sign} \sigma)(\operatorname{sign} \tau) \prod_{k=1}^{p} (A_1)_{\sigma_k,k} \prod_{r=1}^{q} (A_2)_{\tau_r,r} \\ &= (\det A_1)(\det A_2), \end{aligned}$$

the last line by 9.46 applied to A_1 and to A_2. (The entries of C never appear, since no contributing permutation selects a position with $j \leq p < k$.)

Problem 9C.6. Suppose $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an n-by-n matrix, with v_k denoting the kth column of A. Show that if $(m_1, \ldots, m_n) \in \operatorname{perm} n$, then

$$\det \begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix} = \big(\operatorname{sign}(m_1, \ldots, m_n)\big) \det A.$$

Solution. Define $\alpha\colon (\mathbf{F}^n)^n \to \mathbf{F}$ by

$$\alpha(u_1, \ldots, u_n) = \det \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix};$$

by 9.45 this is an alternating n-linear form on $\mathbf{F}^n$. Applying 9.35 to α with the permutation $(m_1, \ldots, m_n)$ and the columns $v_1, \ldots, v_n$ of A,

$$\det\begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix} = \big(\operatorname{sign}(m_1, \ldots, m_n)\big)\,\alpha(v_1, \ldots, v_n) = \big(\operatorname{sign}(m_1, \ldots, m_n)\big) \det A.$$

Problem 9C.7. Suppose $T \in \mathcal{L}(V)$ is invertible. Let p denote the characteristic polynomial of T and let q denote the characteristic polynomial of T^{-1}. Prove that

$$q(z) = \frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right)$$

for all nonzero $z \in \mathbf{F}$.

Solution. Let $n = \dim V$, so $p(z) = \det(zI - T)$ and $q(z) = \det(zI - T^{-1})$ by 9.63. The third bullet point of 9.42 together with $\det T \neq 0$ (9.50) gives

$$p(0) = \det(-T) = (-1)^n \det T \neq 0, \qquad \frac{1}{p(0)} = \frac{(-1)^n}{\det T}.$$

Now fix $z \in \mathbf{F}$ with $z \neq 0$. Then $zI - T^{-1} = -T^{-1}(I - zT)$ (expand the right side) and $I - zT = z(\frac{1}{z}I - T)$, so multiplicativity of the determinant (9.49(a)) and the third bullet point of 9.42 twice give

$$\begin{aligned} q(z) &= \big(\det(-T^{-1})\big)\big(\det(I - zT)\big) \\ &= \frac{(-1)^n}{\det T} \cdot z^n \det\left(\frac{1}{z}I - T\right) = \frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right), \end{aligned}$$

using $\det(-T^{-1}) = (-1)^n \det(T^{-1}) = (-1)^n / \det T$ (9.50) and the definition of p.

Problem 9C.8. Suppose $T \in \mathcal{L}(V)$ is an operator with no eigenvalues (which implies that $\mathbf{F} = \mathbf{R}$). Prove that $\det T > 0$.

Solution. Let $n = \dim V \geq 1$ (here $\mathbf{F} = \mathbf{R}$, by 5.19) and let $p(x) = \det(xI - T)$ be the characteristic polynomial of T (9.63), a monic polynomial of degree n with real coefficients (the observation just after 9.63). By 9.51 the real zeros of p are exactly the eigenvalues of T, so p has none; hence the factorization 4.16 of the nonconstant p over $\mathbf{R}$ has no linear factors:

$$p(x) = c\,(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M), \qquad b_k^2 < 4c_k.$$

Comparing leading coefficients and degrees gives $c = 1$ and $n = 2M$, so n is even. Each $c_k > b_k^2/4 \geq 0$, so by the third bullet point of 9.42,

$$\det T = (-1)^n \det T = \det(-T) = p(0) = c_1 \cdots c_M > 0.$$

Problem 9C.9. Suppose that V is a real vector space of even dimension, $T \in \mathcal{L}(V)$, and $\det T < 0$. Prove that T has at least two distinct eigenvalues.

Solution. Suppose instead that T has at most one distinct eigenvalue. Let $n = \dim V$, even, and let $p(x) = \det(xI - T)$ be the characteristic polynomial of T (9.63), monic of degree n with real coefficients; by the third bullet point of 9.42 and the evenness of n,

$$p(0) = \det(-T) = (-1)^n \det T = \det T < 0,$$

and $\det T \neq 0$ makes T invertible (9.50), so 0 is not an eigenvalue. Factor the nonconstant p over $\mathbf{R}$ by 4.16 and compare leading coefficients and degrees:

$$p(x) = (x - \lambda)^m q(x), \qquad q(x) = (x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M),$$

with $b_k^2 < 4c_k$ and $m + 2M = n$; here all the linear factors share the same root λ, since completing the square shows each quadratic factor has no real zero, so by 9.51 the eigenvalues of T are exactly the λ_j occurring and by hypothesis there is at most one such value. Each $c_k > b_k^2/4 \geq 0$, so $q(0) = c_1 \cdots c_M > 0$, and $m = n - 2M$ is even.
(i) $m = 0$: then $p(0) = q(0) > 0$.
(ii) $m > 0$: then λ is an eigenvalue, so $\lambda \neq 0$, and $p(0) = (-\lambda)^m q(0) = \lambda^m q(0) > 0$.
Either way $p(0) > 0$, contradicting $p(0) < 0$. Hence T has at least two distinct eigenvalues.

Problem 9C.10. Suppose V is a real vector space of odd dimension and $T \in \mathcal{L}(V)$. Without using the minimal polynomial, prove that T has an eigenvalue.
[This result was previously proved without using determinants or the characteristic polynomial; see 5.34.]

Solution. Let $n = \dim V$, odd, and let $p(x) = \det(xI - T)$ be the characteristic polynomial of T (9.63). Applying 9.46 to $xI - A$, where $A = \mathcal{M}(T)$ with respect to any basis (9.53), the identity permutation contributes $(x - A_{1,1}) \cdots (x - A_{n,n})$ while every other permutation moves at least two indices and so contributes degree at most $n - 2$; hence p is monic of degree n with real coefficients (no minimal polynomial is used). By 4.16 there are real $c, \lambda_1, \ldots, \lambda_m, b_1, \ldots, b_M, c_1, \ldots, c_M$ with

$$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m) \cdot (x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M),$$

so comparing degrees gives $n = m + 2M$; since n is odd, m is odd and in particular $m \geq 1$. Thus $\det(\lambda_1 I - T) = p(\lambda_1) = 0$, which by 9.51 says that λ_1 is an eigenvalue of T.

Problem 9C.11. Prove or give a counterexample: If $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\det T > 0$, then T has a square root.
[If $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $\det T \neq 0$, then T has a square root (see 8.41).]

Solution. False: take $V = \mathbf{R}^2$ and $T(x, y) = (-x, -2y)$, whose matrix with respect to the standard basis is diagonal with $-1, -2$ on the diagonal, so $\det T = 2 > 0$ by 9.53 and 9.48.
Suppose $S \in \mathcal{L}(\mathbf{R}^2)$ satisfied $S^2 = T$. Then $ST = SS^2 = S^2 S = TS$, so $E(-1, T)$ is invariant under S: if $Tv = -v$, then $T(Sv) = S(Tv) = -(Sv)$. Since $(T+I)(x, y) = (0, -y)$, we have $E(-1, T) = \operatorname{span}(e_1)$, so $Se_1 = ae_1$ for some $a \in \mathbf{R}$. Applying S again,

$$a^2 e_1 = S^2 e_1 = Te_1 = -e_1,$$

so $a^2 = -1$, impossible in $\mathbf{R}$.

Problem 9C.12. Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Define $p\colon \mathbf{F} \to \mathbf{F}$ by

$$p(z) = \det(zS - T).$$

Prove that p is a polynomial of degree $\dim V$ and that the coefficient of $z^{\dim V}$ in this polynomial is $\det S$.

Solution. Because S is invertible, $zS - T = S(zI - S^{-1}T)$ for every $z \in \mathbf{F}$, so multiplicativity of the determinant (9.49(a)) gives

$$p(z) = (\det S)\, q(z), \qquad \text{where } q(z) = \det(zI - S^{-1}T).$$

Here q is the characteristic polynomial of $S^{-1}T$ (9.63), and it is monic of degree $n = \dim V$: applying 9.46 to $zI - A$ with $A = \mathcal{M}(S^{-1}T)$ (9.53), the identity permutation contributes the monic degree-n product $(z - A_{1,1}) \cdots (z - A_{n,n})$, while every other permutation moves at least two indices and so contributes degree at most $n - 2$. Since $\det S \neq 0$ (9.50), the polynomial $p = (\det S)q$ has degree exactly $\dim V$ and its coefficient of $z^{\dim V}$ is $\det S$.

Problem 9C.13. Suppose $\mathbf{F} = \mathbf{C}$, $T \in \mathcal{L}(V)$, and $n = \dim V > 2$. Let $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.
(a) Find a formula for the coefficient of z^{n-2} in the characteristic polynomial of T in terms of $\lambda_1, \dots, \lambda_n$.
(b) Find a formula for the coefficient of z in the characteristic polynomial of T in terms of $\lambda_1, \dots, \lambda_n$.

Solution. By 9.62, the characteristic polynomial of T is the product of $(z - \mu_j)^{d_j}$ over the distinct eigenvalues μ_j with multiplicities d_j; since $\mathbf{F} = \mathbf{C}$ those multiplicities sum to n (8.25), so that product is

$$q(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n).$$

Expanding by choosing $-\lambda_k$ from the factors indexed by a set $A \subseteq \{1, \dots, n\}$ and z from the rest,

$$q(z) = \sum_{r=0}^{n} (-1)^r e_r z^{n-r}, \qquad e_r = \sum_{1 \le k_1 < \cdots < k_r \le n} \lambda_{k_1} \cdots \lambda_{k_r}.$$

(a) Take $r = 2$: the coefficient of z^{n-2} is

$$e_2 = \sum_{1 \le j < k \le n} \lambda_j \lambda_k.$$

(b) The coefficient of $z = z^1$ is the case $r = n - 1$, and a set of size $n - 1$ is the complement of one index, so it equals

$$(-1)^{n-1} e_{n-1} = (-1)^{n-1} \sum_{k=1}^{n} \prod_{j \ne k} \lambda_j.$$

Problem 9C.14. Suppose V is an inner product space and T is a positive operator on V. Prove that

$$\det \sqrt{T} = \sqrt{\det T}.$$

Solution. Both T and its unique positive square root $\sqrt{T}$ (7.39, 7.40) are positive operators, so $\det T \ge 0$ and $\det \sqrt{T} \ge 0$ by 9.59. Multiplicativity of the determinant (9.49(a)) applied to $\sqrt{T}\sqrt{T} = T$ gives

$$\left(\det \sqrt{T}\right)^2 = \det T.$$

Thus $\det \sqrt{T}$ is the (unique) nonnegative real square root of $\det T$, that is, $\det \sqrt{T} = \sqrt{\det T}$.

Problem 9C.15. Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Use the polar decomposition to give a proof that

$$|\det T| = \sqrt{\det(T^*T)}$$

that is different from the proof given earlier (see 9.60).

Solution. By the polar decomposition 7.93 there is a unitary $S \in \mathcal{L}(V)$ with $T = SR$, where $R = \sqrt{T^*T}$ is the unique positive square root (7.39, 7.40) of the positive operator T^*T (self-adjoint since $(T^*T)^* = T^*T$, and $\langle T^*Tv, v\rangle = \|Tv\|^2 \ge 0$). Multiplicativity of the determinant (9.49(a)) together with $|\det S| = 1$ (9.58) and $\det R \ge 0$ (9.59) gives

$$|\det T| = |\det S|\,|\det R| = |\det R| = \det R,$$

while 9.49(a) applied to $R^2 = T^*T$ gives

$$(\det R)^2 = \det(T^*T).$$

Since $\det R$ and $\det(T^*T)$ are both nonnegative (9.59), taking nonnegative square roots yields $|\det T| = \det R = \sqrt{\det(T^*T)}$.

Problem 9C.16. Suppose $T \in \mathcal{L}(V)$. Define $g\colon \mathbf{F} \to \mathbf{F}$ by $g(x) = \det(I + xT)$. Show that $g'(0) = \operatorname{tr} T$.
[Look for a clean solution to this exercise, without using the explicit but complicated formula for the determinant of a matrix.]

Solution. Fix a basis of V, let $A = \mathcal{M}(T)$ with respect to it (so $\operatorname{tr} T = A_{1,1} + \cdots + A_{n,n}$, where $n = \dim V$), write $A_{\cdot,k}$ for the k^{th} column of A, and put

$$D(v_1, \ldots, v_n) = \det \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix},$$

an alternating n-linear form on $\mathbf{F}^n$ by 9.45. Since $\mathcal{M}(I+xT) = I + xA$, whose k^{th} column is $e_k + xA_{\cdot,k}$, 9.53 and linearity in each of the n slots give

$$g(x) = D(e_1 + xA_{\cdot,1}, \ldots, e_n + xA_{\cdot,n}) = \sum_{m=0}^{n} c_m x^m,$$

where c_m sums $D(u_1^S, \ldots, u_n^S)$ over the m-element sets $S \subseteq \{1, \ldots, n\}$, with $u_k^S = A_{\cdot,k}$ for $k \in S$ and $u_k^S = e_k$ otherwise. So g is a polynomial function and $g'(0) = c_1$. Expanding $A_{\cdot,k} = \sum_j A_{j,k} e_j$ in the one displaced slot and using that D is alternating (the terms $j \neq k$ repeat a vector, hence vanish),

$$c_1 = \sum_{k=1}^{n} D(e_1, \ldots, e_{k-1}, A_{\cdot,k}, e_{k+1}, \ldots, e_n) = \sum_{k=1}^{n} A_{k,k} \det I = \operatorname{tr} T.$$

Problem 9C.17. Suppose a, b, c are positive numbers. Find the volume of the ellipsoid

$$\left\{(x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}$$

by finding a set $\Omega \subseteq \mathbf{R}^3$ whose volume you know and an operator T on $\mathbf{R}^3$ such that $T(\Omega)$ equals the ellipsoid above.

Solution. The ellipsoid E equals $T(\Omega)$, where Ω is the open unit ball of $\mathbf{R}^3$, of volume $\frac{4\pi}{3}$, and $T(x, y, z) = (ax, by, cz)$: indeed

$$\frac{(ax)^2}{a^2} + \frac{(by)^2}{b^2} + \frac{(cz)^2}{c^2} = x^2 + y^2 + z^2,$$

and T is a bijection of $\mathbf{R}^3$, with $T^{-1}(u, v, w) = (u/a, v/b, w/c)$, so the two conditions correspond. The matrix of T with respect to the standard basis is diagonal with a, b, c on the diagonal, so $\det T = abc$ by 9.53 and 9.48. Since $a, b, c > 0$, the volume-change formula 9.61 gives

$$\operatorname{volume} E = |\det T| \, (\operatorname{volume} \Omega) = abc \cdot \frac{4\pi}{3} = \frac{4\pi abc}{3}.$$

Problem 9C.18. Suppose that A is an invertible square matrix. Prove that Hadamard's inequality (9.66) is an equality if and only if each column of A is orthogonal to the other columns.

Solution. Let $v_1, \ldots, v_n$ be the columns of A. These are linearly independent, so the QR factorization 7.58 gives $A = QR$ with Q unitary and R upper triangular with positive diagonal entries. Since Q is an isometry, $\|v_k\| = \|QR_{\cdot,k}\| = \|R_{\cdot,k}\|$, while 9.49(b), 9.58 and 9.48 give $|\det A| = |\det R| = \prod_k R_{k,k}$. Upper-triangularity gives, factor by factor,

$$\|R_{\cdot,k}\|^2 = R_{k,k}^2 + \sum_{j<k} |R_{j,k}|^2 \geq R_{k,k}^2,$$

so $0 < R_{k,k} \leq \|R_{\cdot,k}\|$, with equality exactly when $R_{j,k} = 0$ for all $j < k$. All factors on both sides being positive, one strict factor makes the product strict; hence

$$|\det A| = \prod_{k=1}^{n} \|v_k\| \iff R \text{ is diagonal.}$$

It remains to see that R is diagonal if and only if the columns of A are pairwise orthogonal.
(i) If R is diagonal, then $v_k = QR_{\cdot,k} = R_{k,k}Q_{\cdot,k}$, and the columns of the unitary Q form an orthonormal list, so for $j \neq k$

$$\langle v_j, v_k \rangle = R_{j,j}\, \overline{R_{k,k}}\, \langle Q_{\cdot,j}, Q_{\cdot,k} \rangle = 0.$$

(ii) Conversely, if the v_k are pairwise orthogonal, then $R_{\cdot,k} = Q^* v_k$ and Q^* preserves inner products, so the columns of R are pairwise orthogonal. Induct on k to get $R_{\cdot,k} = R_{k,k}e_k$: this holds for $k = 1$ by upper-triangularity, and if $R_{\cdot,j} = R_{j,j}e_j$ for every $j < k$, then

$$0 = \langle R_{\cdot,k}, R_{\cdot,j} \rangle = \overline{R_{j,j}}\, R_{j,k}$$

with $R_{j,j} > 0$ forces $R_{j,k} = 0$ for every $j < k$, and $R_{j,k} = 0$ for $j > k$ already.

Problem 9C.19. Suppose V is an inner product space, $e_1, \ldots, e_n$ is an orthonormal basis of V, and $T \in \mathcal{L}(V)$ is a positive operator.
(a) Prove that $\det T \leq \prod_{k=1}^{n} \langle Te_k, e_k \rangle$.
(b) Prove that if T is invertible, then the inequality in (a) is an equality if and only if e_k is an eigenvector of T for each $k = 1, \ldots, n$.

Solution. Let $R = \sqrt{T}$ (7.39, 7.40), so $R^*R = R^2 = T$, and let $B = \mathcal{M}(R, (e_1, \ldots, e_n))$, with columns $B_{\cdot,k}$. By 9.49(a), 9.56(c) and 9.53,

$$\det T = \det(R^*R) = \overline{\det R}\, \det R = |\det B|^2,$$

while $Re_k = \sum_j B_{j,k} e_j$ with $e_1, \ldots, e_n$ orthonormal gives, by Parseval's identity 6.30(b),

$$\langle Te_k, e_k \rangle = \langle Re_k, Re_k \rangle = \|Re_k\|^2 = \|B_{\cdot,k}\|^2.$$

(a) Hadamard's inequality 9.66 applied to B gives $|\det B| \leq \prod_k \|B_{\cdot,k}\|$; squaring (both sides are nonnegative),

$$\det T = |\det B|^2 \leq \prod_{k=1}^{n} \|B_{\cdot,k}\|^2 = \prod_{k=1}^{n} \langle Te_k, e_k \rangle.$$

(b) If T is invertible, then $\det T \neq 0$ (9.50), so B is invertible; since $x \mapsto x^2$ is injective on $[0, \infty)$, the inequality in (a) is an equality exactly when Hadamard's inequality for B is, which by Exercise 9C.18 happens exactly when the columns of B are pairwise orthogonal. Coordinates with respect to an orthonormal basis preserve inner products (6.30(c)), so for $j \neq k$

$$\langle B_{\cdot,j}, B_{\cdot,k} \rangle = \langle Re_j, Re_k \rangle = \langle Te_j, e_k \rangle,$$

and since $Te_j = \sum_k \langle Te_j, e_k \rangle e_k$ by 6.30(a), the vanishing of all of these is equivalent to $Te_j = \langle Te_j, e_j \rangle e_j$ for each j, that is, to each e_j being an eigenvector of T.

Problem 9C.20. Suppose A is an n-by-n matrix, and suppose c is such that $|A_{j,k}| \leq c$ for all $j, k \in \{1, \ldots, n\}$. Prove that
$$|\det A| \leq c^n n^{n/2}.$$
[The formula for the determinant of a matrix (9.46) shows that $|\det A| \leq c^n n!$. However, the estimate given by this exercise is much better. For example, if $c = 1$ and $n = 100$, then $c^n n! \approx 10^{158}$, but the estimate given by this exercise is the much smaller number 10^{100}. If n is an integer power of 2, then the inequality above is sharp and cannot be improved.]

Solution. Let $v_1, \ldots, v_n$ be the columns of A; note $c \geq |A_{1,1}| \geq 0$. For each k, the hypothesis gives
$$\|v_k\|^2 = \sum_{j=1}^{n} |A_{j,k}|^2 \leq nc^2,$$
so $\|v_k\| \leq c\sqrt{n}$. Since all these quantities are nonnegative, Hadamard's inequality 9.66 then gives
$$|\det A| \leq \prod_{k=1}^{n} \|v_k\| \leq \left(c\sqrt{n}\right)^n = c^n n^{n/2}.$$

Problem 9C.21. Suppose n is a positive integer and $\delta\colon \mathbf{C}^{n,n} \to \mathbf{C}$ is a function such that
$$\delta(AB) = \delta(A) \cdot \delta(B)$$
for all $A, B \in \mathbf{C}^{n,n}$ and $\delta(A)$ equals the product of the diagonal entries of A for each diagonal matrix $A \in \mathbf{C}^{n,n}$. Prove that
$$\delta(A) = \det A$$
for all $A \in \mathbf{C}^{n,n}$.
[Recall that $\mathbf{C}^{n,n}$ denotes the set of n-by-n matrices with entries in $\mathbf{C}$. This exercise shows that the determinant is the unique function defined on square matrices that is multiplicative and has the desired behavior on diagonal matrices. This result is analogous to Exercise 10 in Section 8D, which shows that the trace is uniquely determined by its algebraic properties.]

Solution. First, $\delta(I) = 1$, since I is diagonal with all diagonal entries 1. Hence if A is invertible then $\delta(A)\delta(A^{-1}) = \delta(I) = 1$, so $\delta(A) \neq 0$, and since complex numbers commute
$$\delta(SAS^{-1}) = \delta(S)\,\delta(A)\,\delta(S^{-1}) = \delta(A)\,\delta(S)\,\delta(S^{-1}) = \delta(A)$$
for every invertible S. Below $E^{(j,k)}$ is the matrix with 1 in row j, column k and 0 elsewhere, and a matrix is identified with its operator on $\mathbf{C}^n$ relative to the standard basis.
(i) A not invertible. Then $\det A = 0$ (9.50). Pick $v \neq 0$ with $Av = 0$, extend to a basis $v_1, \ldots, v_{n-1}, v$ (2.32), and let Q be the (invertible) matrix with those columns; the columns of AQ are $Av_1, \ldots, Av_{n-1}, 0$. Right multiplication by a diagonal D scales column k by $D_{k,k}$, so with D_0 diagonal with entries $1, \ldots, 1, 0$ we get $(AQ)D_0 = AQ$ and hence, since $\delta(D_0) = 0$,
$$\delta(A)\,\delta(Q) = \delta(AQ) = \delta(AQ)\,\delta(D_0) = 0.$$
As $\delta(Q) \neq 0$, this gives $\delta(A) = 0 = \det A$.
(ii) $\delta(I + aE^{(j,k)}) = 1$ for $j \neq k$ and $a \in \mathbf{C}$. Write $T_{j,k}(a) = I + aE^{(j,k)}$; since $E^{(j,k)}E^{(j,k)} = 0$, we get $T_{j,k}(a)T_{j,k}(b) = T_{j,k}(a+b)$, so $T_{j,k}(a)$ is invertible (inverse $T_{j,k}(-a)$) and $T_{j,k}(a)^2 = T_{j,k}(2a)$. Let D be diagonal with $D_{j,j} = 2$ and all other diagonal entries 1; conjugation by D scales the (j,k) entry by $D_{j,j}/D_{k,k} = 2$, so $DT_{j,k}(a)D^{-1} = T_{j,k}(2a) = T_{j,k}(a)^2$ and therefore
$$\delta\big(T_{j,k}(a)\big) = \delta\big(D\,T_{j,k}(a)\,D^{-1}\big) = \delta\big(T_{j,k}(a)\big)^2.$$
Thus $\delta(T_{j,k}(a)) \in \{0, 1\}$, and it is nonzero by invertibility, hence equals 1.

(iii) $\delta(N)=1$ for every upper-triangular N with 1's on the diagonal. Such an N is a product of matrices $I+aE^{(j,k)}$ with $j<k$, by induction on the size m (the case $m=1$ being the empty product): with $r_k=N_{1,k}$ for $k\geq 2$ and N' obtained from N by deleting the first row and column,

$$N=\begin{pmatrix}1 & 0\\ 0 & N'\end{pmatrix}\begin{pmatrix}1 & r\\ 0 & I\end{pmatrix}, \qquad \begin{pmatrix}1 & r\\ 0 & I\end{pmatrix}=\prod_{k=2}^{m}\left(I+r_kE^{(1,k)}\right),$$

the product collapsing because $E^{(1,j)}E^{(1,k)}=0$ for $j,k\geq 2$; and the first factor is the image of N' under the map placing a matrix in the lower-right block below a leading 1, which is multiplicative and carries $I+aE^{(j,k)}$ to $I+aE^{(j+1,k+1)}$, so the induction hypothesis applies to it. Now multiplicativity and (ii) give $\delta(N)=1$.
(iv) A invertible. By 5.47 the operator A has an upper-triangular matrix U with respect to some basis $u_1,\ldots,u_n$ of $\mathbf{C}^n$; with P the invertible matrix whose columns are $u_1,\ldots,u_n$ we get $APe_k=Au_k=PUe_k$ for each k, so $A=PUP^{-1}$. Hence $\delta(A)=\delta(U)$, while 9.49(b) applied to PUP^{-1} and to $PP^{-1}=I$ gives $\det A=\det U$. Since U is invertible, 9.48 and 9.50 give $\det U=U_{1,1}\cdots U_{n,n}\neq 0$, so with D the diagonal matrix with entries $U_{1,1},\ldots,U_{n,n}$ the matrix $N=D^{-1}U$ is upper triangular with 1's on the diagonal, and

$$\delta(A)=\delta(DN)=\delta(D)\,\delta(N)=U_{1,1}\cdots U_{n,n}=\det U=\det A.$$

9.4 Exercises 9D

Problem 9D.1. Suppose $v\in V$ and $w\in W$. Prove that $v\otimes w=0$ if and only if $v=0$ or $w=0$.

Solution. Both directions come from $(v\otimes w)(\varphi,\tau)=\varphi(v)\tau(w)$ (9.71).
(i) If $v=0$ or $w=0$, then $\varphi(v)\tau(w)=0$ for every $\varphi\in V'$ and $\tau\in W'$, so $v\otimes w=0$.
(ii) If $v\neq 0$ and $w\neq 0$, extend each of the linearly independent lists v and w to a basis (2.32) and apply the linear map lemma 3.4 to get $\varphi\in V'$ and $\tau\in W'$ with $\varphi(v)=\tau(w)=1$; then

$$(v\otimes w)(\varphi,\tau)=\varphi(v)\,\tau(w)=1\neq 0,$$

so $v\otimes w\neq 0$.

Problem 9D.2. Give an example of six distinct vectors v_1,v_2,v_3,w_1,w_2,w_3 in $\mathbb{R}^3$ such that

$$v_1\otimes w_1+v_2\otimes w_2+v_3\otimes w_3=0$$

but none of $v_1\otimes w_1$, $v_2\otimes w_2$, $v_3\otimes w_3$ is a scalar multiple of another element of this list.

Solution. Take, with e_1,e_2,e_3 the standard basis of $\mathbb{R}^3$,

$$\begin{aligned} v_1&=e_1=(1,0,0), & w_1&=e_2=(0,1,0),\\ v_2&=2e_1=(2,0,0), & w_2&=e_3=(0,0,1),\\ v_3&=-e_1=(-1,0,0), & w_3&=e_2+2e_3=(0,1,2),\end{aligned}$$

six distinct vectors. Writing $A=e_1\otimes e_2$ and $B=e_1\otimes e_3$, bilinearity (9.73) gives

$$v_1\otimes w_1=A,\qquad v_2\otimes w_2=2B,\qquad v_3\otimes w_3=-A-2B,$$

whose sum is 0. Since e_1 and e_2,e_3 are linearly independent lists, 9.74(a) makes A,B linearly independent in $\mathbb{R}^3\otimes\mathbb{R}^3$. Each of the six statements "one of A, $2B$, $-A-2B$ equals λ times another" rearranges to $\alpha A+\beta B=0$ whose coefficients α,β do not both vanish for any $\lambda\in\mathbb{R}$, hence is impossible. (Check!)

Problem 9D.3. Suppose that $v_1, \dots, v_m$ is a linearly independent list in V. Suppose also that $w_1, \dots, w_m$ is a list in W such that

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

Prove that $w_1 = \cdots = w_m = 0$.

Solution. Fix k. Extend $v_1, \dots, v_m$ to a basis of V (2.32) and use the linear map lemma 3.4 to obtain $\varphi \in V'$ with $\varphi(v_k) = 1$, $\varphi(v_j) = 0$ for $j \neq k$, and $\varphi = 0$ on the added basis vectors. Applying the zero functional $v_1 \otimes w_1 + \cdots + v_m \otimes w_m$ to (φ, τ), where $\tau \in W'$ is arbitrary, gives by 9.71

$$0 = \sum_{j=1}^{m} \varphi(v_j)\,\tau(w_j) = \tau(w_k).$$

Were $w_k \neq 0$, then 2.32 and 3.4 would supply $\tau \in W'$ with $\tau(w_k) = 1$ (as in Exercise 9D.1). Hence $w_k = 0$ for every k.

Problem 9D.4. Suppose $\dim V > 1$ and $\dim W > 1$. Prove that

$$\{v \otimes w : (v, w) \in V \times W\}$$

is not a subspace of $V \otimes W$.
This exercise implies that if $\dim V > 1$ and $\dim W > 1$, then

$$\{v \otimes w : (v, w) \in V \times W\} \neq V \otimes W.$$

Solution. The set $E = \{v \otimes w : (v, w) \in V \times W\}$ is not closed under addition. Choose linearly independent lists v_1, v_2 in V and w_1, w_2 in W (possible since $\dim V > 1$ and $\dim W > 1$), and suppose $v_1 \otimes w_1 + v_2 \otimes w_2 = v \otimes w$ for some $v \in V$ and $w \in W$. Extend v_1, v_2 to a basis $v_1, \dots, v_n$ of V (2.32) and write $v = a_1 v_1 + \cdots + a_n v_n$; bilinearity (9.73) then turns the supposed equation into

$$v_1 \otimes (a_1 w - w_1) + v_2 \otimes (a_2 w - w_2) + \sum_{j=3}^{n} v_j \otimes (a_j w) = 0.$$

Since $v_1, \dots, v_n$ is linearly independent, Exercise 9D.3 forces every second slot to vanish, so $w_1 = a_1 w$ and $w_2 = a_2 w$. Then the linearly independent list w_1, w_2 lies in $\operatorname{span}(w)$, contradicting 2.22. So $v_1 \otimes w_1 + v_2 \otimes w_2$ is not in E, and E is not a subspace of $V \otimes W$.

Problem 9D.5. Suppose m and n are positive integers. For $v \in \mathbb{F}^m$ and $w \in \mathbb{F}^n$, identify $v \otimes w$ with an m-by-n matrix as in Example 9.76. With that identification, show that the set

$$\{v \otimes w : v \in \mathbb{F}^m \text{ and } w \in \mathbb{F}^n\}$$

is the set of m-by-n matrices (with entries in $\mathbb{F}$) that have rank at most one.

Solution. Write $M(v, w)$ for the matrix identified with $v \otimes w$ in 9.76, so $M(v, w)_{j,k} = v_j w_k$, and write $\hat{u} \in \mathbb{F}^{m,1}$ for the column vector corresponding to $u \in \mathbb{F}^m$; the rank of a matrix is the dimension of the span of its columns (3.58 and 3.52). The two sets coincide.
(i) $\subseteq$: column k of $M(v, w)$ is $w_k \hat{v}$, so the span of the columns lies in $\operatorname{span}(\hat{v})$, whence $\operatorname{rank} M(v, w) \leq 1$ by 2.37.
(ii) $\supseteq$: suppose $\operatorname{rank} A \leq 1$. If $\operatorname{rank} A = 0$, then every column of A is 0, so $A = 0 = M(0, 0)$. Otherwise the span of the columns is one-dimensional; let $\hat{v}$ be a basis of it, so for each k there is $w_k \in \mathbb{F}$ with column k of A equal to $w_k \hat{v}$. Comparing entries in row j gives $A_{j,k} = v_j w_k$, so $A = M(v, w)$ with

$w = (w_1, \ldots, w_n)$.

Problem 9D.6. Suppose m and n are positive integers. Give a description, analogous to Exercise 5, of the set of m-by-n matrices (with entries in $\mathbb{F}$) that have rank at most two.

Solution. With the identification of 9.76, the m-by-n matrices of rank at most two are exactly the sums of two matrices of rank at most one:

$$\{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in \mathbb{F}^m \text{ and } w_1, w_2 \in \mathbb{F}^n\}.$$

Keep the notation $M(v, w)$ and $\hat{u}$ of Exercise 9D.5.
(i) $\subseteq$: column k of $M(v_1, w_1) + M(v_2, w_2)$ is $(w_1)_k \hat{v_1} + (w_2)_k \hat{v_2}$, so the span of the columns lies in $\operatorname{span}(\hat{v_1}, \hat{v_2})$, of dimension at most 2 by 2.22; hence the rank is at most 2 by 2.37.
(ii) $\supseteq$: suppose $\operatorname{rank} A \leq 2$ and let C be the span of the columns of A. Take a basis of C, padded with zero vectors if necessary, to get a list $\hat{v_1}, \hat{v_2}$ with $\operatorname{span}(\hat{v_1}, \hat{v_2}) = C$. For each k write column k of A as $c_k \hat{v_1} + d_k \hat{v_2}$; comparing entries in row j gives $A_{j,k} = (v_1)_j c_k + (v_2)_j d_k$, so $A = M(v_1, w_1) + M(v_2, w_2)$ with $w_1 = (c_1, \ldots, c_n)$ and $w_2 = (d_1, \ldots, d_n)$.

Problem 9D.7. Suppose $\dim V > 2$ and $\dim W > 2$. Prove that

$$\{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in V \text{ and } w_1, w_2 \in W\} \neq V \otimes W.$$

Solution. Pick linearly independent lists x_1, x_2, x_3 in V and y_1, y_2, y_3 in W (possible because $\dim V > 2$ and $\dim W > 2$); then

$$z = x_1 \otimes y_1 + x_2 \otimes y_2 + x_3 \otimes y_3 \in V \otimes W$$

lies outside the set T on the left side.
Indeed, suppose $z = v_1 \otimes w_1 + v_2 \otimes w_2$. Put $U = \operatorname{span}(v_1, v_2)$, so $r = \dim U \leq 2$ by 2.22; take a basis $u_1, \ldots, u_r$ of U and extend it to a basis $u_1, \ldots, u_n$ of V by 2.32. Writing $v_i = \sum_{p=1}^{r} b_{i,p} u_p$ and $x_l = \sum_{p=1}^{n} \lambda_{l,p} u_p$, bilinearity of the tensor product (9.73) gives

$$\begin{aligned} \sum_{p=1}^{n} u_p \otimes c_p &= v_1 \otimes w_1 + v_2 \otimes w_2 \\ &= z = \sum_{p=1}^{n} u_p \otimes d_p, \end{aligned}$$

where $c_p = b_{1,p} w_1 + b_{2,p} w_2$ for $p \leq r$ while $c_p = 0$ for $p > r$, and $d_p = \lambda_{1,p} y_1 + \lambda_{2,p} y_2 + \lambda_{3,p} y_3$. Because $u_1, \ldots, u_n$ is linearly independent, Exercise 9D.3 applied to $\sum_{p=1}^{n} u_p \otimes (d_p - c_p) = 0$ yields $d_p = c_p$ for every p. Hence $d_p = 0$ for $p > r$, so the linear independence of y_1, y_2, y_3 forces $\lambda_{1,p} = \lambda_{2,p} = \lambda_{3,p} = 0$ for every $p > r$. Thus x_1, x_2, x_3 is a linearly independent list of length 3 in $U = \operatorname{span}(u_1, \ldots, u_r)$, contradicting $r \leq 2$ via 2.22.

Problem 9D.8. Suppose $v_1, \ldots, v_m \in V$ and $w_1, \ldots, w_m \in W$ are such that

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

Suppose that U is a vector space and $\Gamma\colon V \times W \to U$ is a bilinear map. Show that

$$\Gamma(v_1, w_1) + \cdots + \Gamma(v_m, w_m) = 0.$$

Solution. Apply the linear map $\hat{\Gamma}\colon V \otimes W \to U$ that 9.79(a) attaches to the bilinear map Γ, which satisfies $\hat{\Gamma}(v \otimes w) = \Gamma(v, w)$ for all $(v, w) \in V \times W$. Since $\hat{\Gamma}$ is linear it sends 0 to 0 (3.10) and preserves sums, so

$$\begin{aligned} 0 &= \hat{\Gamma}(v_1 \otimes w_1 + \cdots + v_m \otimes w_m) \\ &= \hat{\Gamma}(v_1 \otimes w_1) + \cdots + \hat{\Gamma}(v_m \otimes w_m) \\ &= \Gamma(v_1, w_1) + \cdots + \Gamma(v_m, w_m). \end{aligned}$$

Problem 9D.9. Suppose $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Prove that there exists a unique operator on $V \otimes W$ that takes $v \otimes w$ to $Sv \otimes Tw$ for all $v \in V$ and $w \in W$.
[In an abuse of notation, the operator on $V \otimes W$ given by this exercise is often called $S \otimes T$.]

Solution. The operator is the map $\hat{\Gamma}$ attached by 9.79(a) to the map $\Gamma\colon V \times W \to V \otimes W$ defined by $\Gamma(v, w) = Sv \otimes Tw$.
This Γ is bilinear (9.77): for fixed $w \in W$, the linearity of S and the bilinearity of the tensor product (9.73) give

$$\begin{aligned} \Gamma(\lambda v_1 + v_2, w) &= (\lambda S v_1 + S v_2) \otimes Tw \\ &= \lambda \Gamma(v_1, w) + \Gamma(v_2, w), \end{aligned}$$

and the same computation in the other slot, using the linearity of T and 9.73, shows $w \mapsto \Gamma(v, w)$ is linear for each fixed $v \in V$. Hence 9.79(a) supplies a unique linear map $\hat{\Gamma}\colon V \otimes W \to V \otimes W$ with

$$\hat{\Gamma}(v \otimes w) = \Gamma(v, w) = Sv \otimes Tw$$

for all $v \in V$ and $w \in W$. It maps $V \otimes W$ to itself, so it is an operator on $V \otimes W$, and the uniqueness clause of 9.79(a) is exactly the asserted uniqueness.

Problem 9D.10. Suppose $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Prove that $S \otimes T$ is an invertible operator on $V \otimes W$ if and only if both S and T are invertible operators. Also, prove that if both S and T are invertible operators, then $(S \otimes T)^{-1} = S^{-1} \otimes T^{-1}$, where we are using the notation from the comment after Exercise 9.

Solution. Two operators on $V \otimes W$ that agree on every elementary tensor agree on the basis $\{e_j \otimes f_k\}$ of 9.74(b) and so are equal; this converts the computation

$$\begin{aligned} (S \otimes T)\big((S' \otimes T')(v \otimes w)\big) &= SS'v \otimes TT'w \\ &= \big((SS') \otimes (TT')\big)(v \otimes w) \end{aligned}$$

into the identity $(S \otimes T)(S' \otimes T') = (SS') \otimes (TT')$ for all $S, S' \in \mathcal{L}(V)$ and $T, T' \in \mathcal{L}(W)$, and likewise gives $I_V \otimes I_W = I_{V \otimes W}$.
(i) If S and T are invertible, then

$$(S \otimes T)(S^{-1} \otimes T^{-1}) = I_V \otimes I_W = I_{V \otimes W}$$

and the same in the other order, so $S \otimes T$ is invertible with $(S \otimes T)^{-1} = S^{-1} \otimes T^{-1}$ (3.59 and 3.60).
(ii) If S is not invertible, then S is not injective (3.65, V being finite-dimensional), so $Sv = 0$ for some $v \neq 0$; choose $w \neq 0$ in W, which is nonzero by the standing assumptions of Chapter 9. Then

$$(S \otimes T)(v \otimes w) = 0 \otimes Tw = 0$$

since $0 \otimes u = 0(0 \otimes u) = 0$ by 9.73, while $v \otimes w \neq 0$ by Exercise 9D.1. So $S \otimes T$ is not injective, hence not invertible by 3.65. Interchanging the roles of the two factors gives the same conclusion when T is not invertible.

Problem 9D.11. Suppose V and W are inner product spaces. Prove that if $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$, then $(S \otimes T)^* = S^* \otimes T^*$, where we are using the notation from the comment after Exercise

9.

Solution. The operator $S^* \otimes T^*$ satisfies the adjoint identity on elementary tensors: for all $v, u \in V$ and $w, x \in W$,

$$\begin{aligned}\langle (S \otimes T)(v \otimes w), u \otimes x \rangle &= \langle Sv, u \rangle \langle Tw, x \rangle \\ &= \langle v, S^* u \rangle \langle w, T^* x \rangle \\ &= \langle v \otimes w, (S^* \otimes T^*)(u \otimes x) \rangle,\end{aligned}$$

where the outer equalities use the description $\langle v \otimes w, u \otimes x \rangle = \langle v, u \rangle \langle w, x \rangle$ of the inner product on $V \otimes W$ (9.82) along with the definitions of $S \otimes T$ and $S^* \otimes T^*$ from Exercise 9D.9, and the middle equality is the defining property 7.1 of the adjoints S^* and T^*.
Each side of the display is linear in the first argument and additive and conjugate homogeneous in the second (6.6), and the elementary tensors include the basis $\{e_j \otimes f_k\}$ of $V \otimes W$ given by 9.74(b); expanding α and β in that basis therefore propagates the identity to

$$\langle (S \otimes T)\alpha, \beta \rangle = \langle \alpha, (S^* \otimes T^*)\beta \rangle$$

for all $\alpha, \beta \in V \otimes W$ (note $V \otimes W$ is finite-dimensional by 9.72, so $(S \otimes T)^*$ exists). Comparing with $\langle (S \otimes T)\alpha, \beta \rangle = \langle \alpha, (S \otimes T)^* \beta \rangle$ from 7.1 gives

$$\langle \alpha, (S \otimes T)^* \beta - (S^* \otimes T^*)\beta \rangle = 0$$

for all α; taking α to be that difference makes its norm 0, so $(S \otimes T)^* \beta = (S^* \otimes T^*)\beta$ for every β.

Problem 9D.12. Suppose that $V_1, \ldots, V_m$ are finite-dimensional inner product spaces. Prove that there is a unique inner product on $V_1 \otimes \cdots \otimes V_m$ such that

$$\langle v_1 \otimes \cdots \otimes v_m,\ u_1 \otimes \cdots \otimes u_m \rangle = \langle v_1, u_1 \rangle \cdots \langle v_m, u_m \rangle$$

for all $(v_1, \ldots, v_m)$ and $(u_1, \ldots, u_m)$ in $V_1 \times \cdots \times V_m$.
[Note that the equation above implies that

$$\|v_1 \otimes \cdots \otimes v_m\| = \|v_1\| \times \cdots \times \|v_m\|$$

for all $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$.]

Solution. The inner product is the one making the basis $\{E_J\}$ below orthonormal. For each k let $n_k = \dim V_k$ and let $e^k_1, \ldots, e^k_{n_k}$ be an orthonormal basis of V_k (6.35); write $J = (j_1, \ldots, j_m)$ for a multi-index with $j_k \in \{1, \ldots, n_k\}$, let $\mathcal{J}$ be the set of all such multi-indices, and set

$$E_J = e^1_{j_1} \otimes \cdots \otimes e^m_{j_m}.$$

By 9.90 the list $\{E_J\}_{J \in \mathcal{J}}$ is a basis of $V_1 \otimes \cdots \otimes V_m$, so

$$\Big\langle \sum_{J \in \mathcal{J}} b_J E_J,\ \sum_{J \in \mathcal{J}} c_J E_J \Big\rangle = \sum_{J \in \mathcal{J}} b_J \overline{c_J}$$

is a well-defined inner product, being the standard inner product on $\mathbf{F}^{n_1 \cdots n_m}$ transported through the coordinate isomorphism of that basis (Check!).
It satisfies the required identity. Definition 9.88 gives

$$(v_1 \otimes \cdots \otimes v_m)(\varphi_1, \ldots, \varphi_m) = \varphi_1(v_1) \cdots \varphi_m(v_m),$$

whose right side is linear in each v_k with the other slots and all $\varphi_i \in V_i'$ fixed; hence $(v_1, \ldots, v_m) \mapsto v_1 \otimes \cdots \otimes v_m$ is m-linear, the m-fold form of 9.73. So writing $v_k = \sum_j a^k_j e^k_j$ and $u_k = \sum_j d^k_j e^k_j$ and expanding each slot in turn, $v_1 \otimes \cdots \otimes v_m = \sum_J a^1_{j_1} \cdots a^m_{j_m} E_J$ and likewise for the u_k, whence

$$\begin{aligned}\langle v_1 \otimes \cdots \otimes v_m,\ u_1 \otimes \cdots \otimes u_m \rangle &= \sum_{J \in \mathcal{J}} \prod_{k=1}^{m} a^k_{j_k} \overline{d^k_{j_k}} \\ &= \prod_{k=1}^{m} \Big(\sum_{j=1}^{n_k} a^k_j \overline{d^k_j} \Big) \\ &= \langle v_1, u_1 \rangle \cdots \langle v_m, u_m \rangle,\end{aligned}$$

the second equality because expanding the product of the m sums produces exactly one term for each multi-index J, and the third because each $e^k_1, \ldots, e^k_{n_k}$ is orthonormal, so that $\langle v_k, u_k \rangle = \sum_j a^k_j \overline{d^k_j}$ by 6.6.

For uniqueness, two such inner products agree at every pair (E_J, E_K) of elementary tensors, and every inner product is linear in its first slot and additive and conjugate homogeneous in its second (6.6), so each of them sends $\alpha = \sum_J b_J E_J$ and $\beta = \sum_K c_K E_K$ to $\sum_J \sum_K b_J \overline{c_K} \langle E_J, E_K \rangle$; hence they coincide. Taking $u_k = v_k$ for each k gives $\|v_1 \otimes \cdots \otimes v_m\|^2 = \|v_1\|^2 \cdots \|v_m\|^2$, the norm identity of the note.

Problem 9D.13. Suppose that $V_1, \ldots, V_m$ are finite-dimensional inner product spaces and $V_1 \otimes \cdots \otimes V_m$ is made into an inner product space using the inner product from Exercise 12. Suppose $e^k_1, \ldots, e^k_{n_k}$ is an orthonormal basis of V_k for each $k = 1, \ldots, m$. Show that the list

$$\{e^1_{j_1} \otimes \cdots \otimes e^m_{j_m}\}_{j_1 = 1, \ldots, n_1;\ \cdots;\ j_m = 1, \ldots, n_m}$$

is an orthonormal basis of $V_1 \otimes \cdots \otimes V_m$.

Solution. Write $J = (j_1, \ldots, j_m)$ and $K = (l_1, \ldots, l_m)$ for multi-indices and put $E_J = e^1_{j_1} \otimes \cdots \otimes e^m_{j_m}$. The defining identity of the inner product from Exercise 9D.12 gives

$$\langle E_J, E_K \rangle = \langle e^1_{j_1}, e^1_{l_1} \rangle \cdots \langle e^m_{j_m}, e^m_{l_m} \rangle,$$

where each factor equals 1 if $j_k = l_k$ and 0 otherwise, since $e^k_1, \ldots, e^k_{n_k}$ is orthonormal. So the product equals 1 when $J = K$ and equals 0 when $J \neq K$, some factor then vanishing; thus $\{E_J\}$ is an orthonormal list. It is also a basis of $V_1 \otimes \cdots \otimes V_m$ by 9.90, each $e^k_1, \ldots, e^k_{n_k}$ being a basis of V_k, and hence an orthonormal basis.

Method (2): the orthonormal list $\{E_J\}$ is linearly independent by 6.25, and its length $n_1 \cdots n_m = (\dim V_1) \cdots (\dim V_m)$ equals $\dim(V_1 \otimes \cdots \otimes V_m)$ by 9.89, so it is a basis by 2.38.